

Copyright (c) 1993 - 2004 Compugen Ltd.

GenCore version 5.1.6

OM protein - protein search, using sw model
 Run on: July 15, 2004, 20:30:05 ; Search time 8.84706 Seconds
 (without alignments)
 276.623 Million cell updates/sec

Title: US-10-004-381-25
 Perfect score: 258

Sequence: 1 MDEKITGWRGGHVVVEGLAGE.....LEHHPOQORPMMGGCKLG 47

Scoring table: BLOSUM62
 Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
 Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
 Listing first 45 summaries

Database : SwissProt_42; *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60.5	23.4	177	1	HSLV_RHIZO
2	59.5	23.1	342	1	YDEY_ECOLI
3	57.5	22.3	1065	1	KDGI_HUMAN
4	57	22.1	474	1	TRPC_CORGII
5	56.5	21.9	1	GPMB_SALTM	
6	56.5	21.9	215	1	GFMF_SALTY
7	56.5	21.9	404	1	KVB3_HUMAN
8	56.5	21.9	415	1	B2V0IN
9	56.5	21.9	1560	1	SMCX_HUMAN
10	56	21.7	261	1	YEV6_YEAST
11	56	21.7	299	1	YJ52_STREPT
12	55.5	21.5	188	1	HSLV_CAUCR
13	55.5	21.5	288	1	AOX4_AERPE
14	55.5	21.5	465	1	FYX3_MOUSE
15	54	20.9	376	1	ALC_STRCO
16	53.5	20.7	352	1	GCP_TREPA
17	53	20.5	404	1	KVB3_RAT
18	53	20.5	348	1	PLSX_SYNTH
19	53	20.5	358	1	AERB_AERPE
20	53	20.5	381	1	PTX3_MOUSE
21	53	20.5	493	1	MURE_CHLKV
22	53	20.5	528	1	WR42_ARATH
23	53	20.5	539	1	ASCR1_ASCR1
24	52.5	20.3	184	1	HSLV_BRUME
25	52.5	20.3	184	1	HSLV_BRUSL
26	52.5	20.3	344	1	PLSX_SYNTH
27	52.5	20.3	461	1	SX10_CHICK
28	52.5	20.3	591	1	VATA_STRP3
29	52	20.2	113	1	XRE_BACSL
30	52	20.2	270	1	BDNF_CYPCA
31	52	20.2	356	1	GBB2_CAEEL
32	52	20.2	378	1	HMAN_DROME
33	52	20.2	449	1	AROF_STRCO

ALIGNMENTS									
RESULT 1	HSLV_RHIZO	STANDARD:	PRT:	177 AA.					
ID	HSLV_RHIZO	AC: 098C7B;	DT: 28-FEB-2003	(Rel. 41, Created)					
		DB: AFP-dependent protease hslv (EC 3.4.25.-).	DT: 10-OCT-2003	(Rel. 42, Last annotation update)					
			GN: HSLV OR MU5007.						
			OS: Rhizobium loti (Mesorhizobium loti).						
			OC: Bacteria; Proteobacteria; Alpha-proteobacteria; Rhizobiales;						
			OC: Phyllobacteriaceae; Mesorhizobium.						
			OX: NCBI_TaxID:381;						
			RN: [1]						
			RP: SEQUENCE FROM N.A.						
			RC: STRAIN-MAP#30309;						
			RX: MEDLINE#21082930; PubMed#11214968;						
			RA: Kaneko T., Nakamura Y., Sato S., Aranizu E., Kato T., Sasamoto S.,						
			RA: Kishida Y., Kiyokawa C., Ishikawa A., Kawashima K., Kimura T.,						
			RA: Matsubara A., Idzuka S., Nakazaki N., Shimpo S., Sugimoto M.,						
			RA: Mochizuki Y., Nakada M., Tabata S.,						
			RA: Takeuchi C., Yamada M., Kohara M., Matsumoto M., Matsuno A.,						
			RA: Watanabe A., Nakamura K., Kohara M., Kawashima K., Kimura T.,						
			RA: Kishida Y., Kiyokawa C., Ishikawa A., Nakazaki N., Shimpo S., Sugimoto M.,						
			RT: "Complete genome structure of the nitrogen-fixing symbiotic bacterium Mesorhizobium loti." [2000]						
			RL: DNA Res. 7, 331-338 (2000).						
			CC: -I- FUNCTION: Protease subunit of a proteaseome-like degradation complex (By similarity).						
			CC: -I- SUBUNIT: A double ring-shaped homohexamer of hslv is capped on each side by a ring-shaped hslv homohexamer (By similarity).						
			CC: -I- SUBCELLULAR LOCATION: Cyttoplasmic (By similarity).						
			CC: -I- SIMILARITY: Belongs to peptidase family TB. Hslv subfamily.						
			CC: This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or_send_an_email_to_licensee@isb-sib.ch).						
			CC: DR: MEROPS; T01_006; -. DR: EMBL; AP03005; BAB5153.1; -. DR: P0078; T01_006; -. DR: P03368; treponema_p033494; rattus_norv_P37950; synchocyst_P37950; aeropyrum_p031060; mus_musculus_P07004; streptomyce_Q823m2; chlamyphili_Q823m2; arachidopsis_Q8xeC3; ascaris_ssu_P23731; brucella_me_Q8ve31; brucella_su_Q8Yf11; brucella_su_Q8dk15; synchococc_Q8w757; gallus_gallus_Q8k8T1; streptococc_P23789; bacillus_su_P23789; cyprinus_ca_Q80323; caenorhabdi_Q80636; drosophila_Q8P0574; streptomyce_P80574						
			DR: Interpro; IPR001353; Peptidase_T1. DR: Pfam; PF00227; proteasome_1. DR: Hydrolase; protease; Threonine protease; Complete proteome. FT: ACT SITE: 4 BY SIMILARITY. SQ: SEQUENCE: 17 AA; 18986 MW; 62F3D3DC6ACB5F05 CRC64;						
			QY: Query Match: 23.4%; Score: 60.5; DB: 1; Length: 177; Best Local Similarity: 41.9%; Pred. No.: 1.7; Matches: 13; Conservative: 6; Mismatches: 5; Indels: 7; Gaps: 1;						
			QY: 9 RGGHVVVEGLAGE-----LQLARLRRHHP 32						
			QY: 42 KCGNNTAGFAGATADFTLRLERAKLEQYP 72						

RESULT 2	SEQUENCE	342 AA;	36394 MW;	419E505026ABDE33	CRC64;
YDEY_ECOLI	STANDARD;	PRT;	342 AA.		
AC	P77672;				
DT	15-JUL-1998 (Rel. 36, Created)				
DT	15-JUN-1998 (Rel. 36, Last sequence update)				
DT	28-FEB-2003 (Rel. 41, Last annotation update)				
DE	Hypothetical ABC transporter permease protein ydeY.				
GN	YDEY OR B1514.				
OS	Escherichia coli.				
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;				
OC	Enterobacteriaceae; Escherichia.				
OX	NCBI_TaxID=562;				
RN	[1] SEQUENCE FROM N.A.				
RP	STRAIN=K12;				
RX	MEDLINE=97251357; PubMed=9097039;				
RA	Aiba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K.,				
RA	Itoh T., Kasai H., Kashimoto K., Kimura S., Kirakawa M.,				
RA	Motomura K., Makino K., Miki T., Mizobuchi K., Mori H., Mori T.,				
RA	Nakade S., Nakamura Y., Nashimoto H., Nishio Y.,				
RA	Oshima T., Saito N., Sampei G., Seki Y., Sivasundaram S.,				
RA	Tagami H., Takeda J., Takeuchi K., Wada C.,				
RA	Yamamoto Y., Horuchi T.;				
RT	"The complete genome sequence of the Escherichia coli K-12 genome corresponding to the 28.0-40.1 min region on the linkage map.",				
RL	DNA Res. 3;363-377(1996)				
CC	--!- FUNCTION: PROBABLY PART OF A BINDING-PROTEIN-DEPENDENT TRANSPORT SYSTEM EGO;YDEY. PROBABLY RESPONSIBLE FOR THE TRANSLOCATION OF THE SUBSTRATE ACROSS THE MEMBRANE.				
CC	--!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane (Potential).				
CC	--!- SIMILARITY: Belongs to the binding-protein-dependent transport system permease family. AraH/bsc subfamily.				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).				
CC	EMBL; AB000249; AAC74587.1; --.				
DR	EMBL; D90794; BAA15201.1; --.				
DR	PIR; E64905; E64905.				
DR	Ecogene; EGI3807; ydeY.				
DR	InterPro; IPR001851; Bac_inmem_transp.				
DR	InterPro; IPR000515; Bpp_transp.				
DR	PF02653; Bpp_transp_2; 1.				
KW	Hypothetical protein; Transport; Transmembrane; Inner membrane; Complete proteome.				
FT	TRANSMEM 14 34 POTENTIAL.				
FT	TRANSMEM 39 59 POTENTIAL.				
FT	TRANSMEM 70 90 POTENTIAL.				
FT	TRANSMEM 93 113 POTENTIAL.				
FT	TRANSMEM 115 135 POTENTIAL.				
FT	TRANSMEM 157 177 POTENTIAL.				
FT	TRANSMEM 213 233 POTENTIAL.				
FT	TRANSMEM 252 272 POTENTIAL.				
FT	TRANSMEM 284 304 POTENTIAL.				
RESULT 3	SEQUENCE	342 AA;	36394 MW;	419E505026ABDE33	CRC64;
KOG1_HUMAN	STANDARD;	PRT;	1065 AA.		
ID	KOG1_HUMAN				
AC	O75912; Q9N249;				
DT	16-OCT-2001 (Rel. 40, Created)				
DT	16-OCT-2001 (Rel. 40, Last sequence update)				
DT	10-OCT-2003 (Rel. 42, Last annotation update)				
DB	diacylglycerol kinase, iota (EC 2.7.1.107) (Diacylglyceride kinase) (DGK-iota) (DAG kinase iota).				
GN	DGK1				
OS	Homo sapiens (Human)				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1] SEQUENCE FROM N.A.				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Retina;				
RX	MEDLINE=9907655; PubMed=9930018;				
RA	Ding L., Traer E., McIntyre T.M., Zimmerman G.A., Prescott S.M.;				
RT	"The cloning and characterization of a novel human diacylglycerol kinase, DGK-iota,"				
RT	Kinase, DGK-iota;"				
RL	J. Biol. Chem. 273:32746-32752(1998).				
RN	[2] SEQUENCE OF 135-1065 FROM N.A. AND VARIANT PHE-153.				
RA	Bowne S.J., Sullivan L.S., Ding L., Traer E., Prescott S.M., Birch D.G., Kennan A., Humphries P., Daiger S.P.;				
RT	Evaluation of human diacylglycerol kinase iota, DGKI, a homolog of Drosophila rdgA, in inherited retinopathy mapping to 7q.;"				
RL	Mol. Vision 6:6-9(2000).				
CC	--!- CATALYTIC ACTIVITY: ATP + 1,2-diacylglycerol = ADP + 1,2-diacylglycerol 3-phosphate.				
CC	--!- SUBCELLULAR LOCATION: Nuclear and cytoplasmic.				
CC	--!- SIMILARITY: Belongs to the eukaryotic diacylglycerol kinase family.				
CC	--!- SIMILARITY: Contains 2 zinc-dependent phorbol-ester and DAG binding domains.				
CC	--!- SIMILARITY: Contains 2 ANK repeats.				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).				
CC	EMBL; AB000249; AAC74587.1; --.				
DR	EMBL; D90794; BAA15201.1; --.				
DR	PIR; E64905; E64905.				
DR	Ecogene; EGI3807; ydeY.				
DR	InterPro; IPR001851; Bac_inmem_transp.				
DR	InterPro; IPR000515; Bpp_transp.				
DR	PF02653; Bpp_transp_2; 1.				
KW	Hypothetical protein; Transport; Transmembrane; Inner membrane; Complete proteome.				
FT	TRANSMEM 14 34 POTENTIAL.				
FT	TRANSMEM 39 59 POTENTIAL.				
FT	TRANSMEM 70 90 POTENTIAL.				
FT	TRANSMEM 93 113 POTENTIAL.				
FT	TRANSMEM 115 135 POTENTIAL.				
FT	TRANSMEM 157 177 POTENTIAL.				
FT	TRANSMEM 213 233 POTENTIAL.				
FT	TRANSMEM 252 272 POTENTIAL.				
FT	TRANSMEM 284 304 POTENTIAL.				
RESULT 4	SEQUENCE	342 AA;	36394 MW;	419E505026ABDE33	CRC64;
YDEY_ECOLI	STANDARD;	PRT;	342 AA.		
ID	YDEY_ECOLI				
AC	P77672;				
DT	15-JUL-1998 (Rel. 36, Created)				
DT	15-JUN-1998 (Rel. 36, Last sequence update)				
DT	28-FEB-2003 (Rel. 41, Last annotation update)				
DE	Hypothetical ABC transporter permease protein ydeY.				
GN	YDEY OR B1514.				
OS	Escherichia coli.				
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;				
OC	Enterobacteriaceae; Escherichia.				
OX	NCBI_TaxID=562;				
RN	[1] SEQUENCE FROM N.A.				
RP	STRAIN=K12;				
RX	MEDLINE=97251357; PubMed=9278503;				
RA	Blattner F.R., Plunkett G., LII, Bloch C.A., Perna N.T., Burland V., Riley M., Collado-Vides J., Glazner J.D., Rose C.K., Mayhew G.F., Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J., Mau B., Shao Y.,				
RA	"The complete genome sequence of Escherichia coli K-12.",				
RA	Science 277:1453-1474(1997).				
RN	[2] SEQUENCE FROM N.A.				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Retina;				
RX	MEDLINE=9907655; PubMed=9930018;				
RA	Ding L., Traer E., McIntyre T.M., Zimmerman G.A., Prescott S.M.;				
RT	"The cloning and characterization of a novel human diacylglycerol kinase, DGK-iota,"				
RT	Kinase, DGK-iota;"				
RL	J. Biol. Chem. 273:32746-32752(1998).				
RN	[2] SEQUENCE OF 135-1065 FROM N.A. AND VARIANT PHE-153.				
RA	Bowne S.J., Sullivan L.S., Ding L., Traer E., Prescott S.M., Birch D.G., Kennan A., Humphries P., Daiger S.P.;				
RT	Evaluation of human diacylglycerol kinase iota, DGKI, a homolog of Drosophila rdgA, in inherited retinopathy mapping to 7q.;"				
RL	Mol. Vision 6:6-9(2000).				
CC	--!- CATALYTIC ACTIVITY: ATP + 1,2-diacylglycerol = ADP + 1,2-diacylglycerol 3-phosphate.				
CC	--!- SUBCELLULAR LOCATION: Nuclear and cytoplasmic.				
CC	--!- SIMILARITY: Belongs to the eukaryotic diacylglycerol kinase family.				
CC	--!- SIMILARITY: Contains 2 zinc-dependent phorbol-ester and DAG binding domains.				
CC	--!- SIMILARITY: Contains 2 ANK repeats.				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).				
CC	EMBL; AB000249; AAC74587.1; --.				
DR	EMBL; D90794; BAA15201.1; --.				
DR	PIR; E64905; E64905.				
DR	Ecogene; EGI3807; ydeY.				
DR	InterPro; IPR001851; Bac_inmem_transp.				
DR	InterPro; IPR000515; Bpp_transp.				
DR	PF02653; Bpp_transp_2; 1.				
KW	Hypothetical protein; Transport; Transmembrane; Inner membrane; Complete proteome.				
FT	TRANSMEM 14 34 POTENTIAL.				
FT	TRANSMEM 39 59 POTENTIAL.				
FT	TRANSMEM 70 90 POTENTIAL.				
FT	TRANSMEM 93 113 POTENTIAL.				
FT	TRANSMEM 115 135 POTENTIAL.				
FT	TRANSMEM 157 177 POTENTIAL.				
FT	TRANSMEM 213 233 POTENTIAL.				
FT	TRANSMEM 252 272 POTENTIAL.				
FT	TRANSMEM 284 304 POTENTIAL.				

DR	EMBL: AF219921; AF43005_1; JOINED.	DE	Tryptophan biosynthesis protein trpcf [Includes: Indole-3-glycerol phosphate synthase (EC 4.1.1.48) (IGPS); N-(5'-phospho-ribosyl)anthranilate isomerase (EC 5.3.1.24) (PR1)].
DR	EMBL: AF219922; AF43005_1; JOINED.	DE	
DR	EMBL: AF219923; AF43006_1; JOINED.	DR	TRPC OR CGC33.
DR	EMBL: AF219924; AF43005_1; JOINED.	OS	Corynebacterium glutamicum (Brevibacterium flumin).
DR	EMBL: AF219925; AF43005_1; JOINED.	OC	Bacteria; Actinobacteria; Actinomycetales;
DR	EMBL: AF219926; AF43006_1; JOINED.	OC	Corynebacterineae; Corynebacteriaceae; Corynebacterium.
DR	EMBL: AF219927; AF43005_1; JOINED.	OX	NCBI_TaxId=1718;
DR	EMBL: AF219928; AF43005_1; JOINED.	RN	[1]
DR	EMBL: AF219929; AF43006_1; JOINED.	RP	SEQUENCE FROM N.A.
DR	EMBL: AF219930; AF43005_1; JOINED.	RX	MEDLINE=87117512; PubMed=3808947;
DR	EMBL: AF219931; AF43005_1; JOINED.	RA	Matsui K., Sano K., Ohtsubo E.;
DR	EMBL: AF219932; AF43006_1; JOINED.	RT	Brevibacterium lactofermentum tryptophan operon.;
DR	EMBL: AF219933; AF43005_1; JOINED.	RT	Brevibacterium lactofermentum tryptophan operon.;
DR	EMBL: AF219934; AF43005_1; JOINED.	RL	Nucleic Acids Res. 14:10113-10114(1986).
DR	EMBL: AF219935; AF43005_1; JOINED.	RN	[2]
DR	EMBL: AF219937; AF43005_1; JOINED.	RP	SEQUENCE FROM N.A.
DR	EMBL: AF219938; AF43006_1; JOINED.	RC	STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
DR	EMBL: AF219939; AF43005_1; JOINED.	RA	Nakagawa S.;
DR	EMBL: AF219940; AF43005_1; JOINED.	RT	"Complete genomic sequence of Corynebacterium glutamicum ATCC 13032.,";
DR	EMBL: AF219941; AF43005_1; JOINED.	RL	Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR	GO: GO:0005634; C-nucleus; TAS.	CC	-!- FUNCTION: Bifunctional enzyme that catalyzes two sequential steps of tryptophan biosynthetic pathway. The first reaction is catalyzed by the isomerase, coded by the trp domain; the second reaction is catalyzed by the synthase, coded by the trc domain.
DR	GO: GO:0004143; F-diacetyl-glycerol kinase activity; TAS.	CC	-!- CATALYTIC ACTIVITY: N-(5-phospho-beta-D-ribofuranosyl)-anthranilate = 1-(2-carboxyphenylamino)-1-deoxy-D-ribulose 5-phosphate.
DR	InterPro: IPR02110; ANK.	CC	-!- CATALYTIC ACTIVITY: 1-(2-carboxyphenylamino)-1-deoxy-D-ribulose 5-phosphate = 1-indol-3-ylglycerol 3-phosphate + CO(2) + H(2)O.
DR	InterPro: IPR00219; DAG PE-bind.	CC	-!- PATHWAY: Tryptophan biosynthesis; third step.
DR	InterPro: IPR00756; DAGKc.	CC	-!- PATHWAY: Tryptophan biosynthesis; fourth step.
DR	InterPro: IPR01206; DAGKc.	CC	-!- SUBUNIT: Monomer.
DR	InterPro: IPR02110; ANK; 2.	CC	-!- SIMILARITY: In the N-terminal section; belongs to the trpc family.
DR	SMART: SM00109; CL; 2.	CC	-!- SIMILARITY: In the C-terminal section; belongs to the trpc family.
DR	SMART: SM00455; DAGKc; 1.	CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
KW	Multigene family; Polymorphism.	CC	
FT	DOMAIN 178 232 PHOROL-ESTER AND DAG BINDING 1.	DR	EMBL: X04960; CAA26626_1; -.
FT	DOMAIN 309 500 PHOROL-ESTER AND DAG BINDING 2.	DR	EMBL: AP005283; BAC0427_1; -.
FT	DOMAIN 374 526 CATALYTIC-A (POTENTIAL).	DR	PIR: B24723; E24723.
FT	DOMAIN 683 990 CATALYTIC-B (POTENTIAL).	DR	HSSP: P01909; IPI.
FT	REPEAT 958 990 ANK 1.	DR	HAMAP: MF_00114; fused; 1.
FT	REPEAT 997 1026 ANK 2.	DR	HAMAP: MF_00115; fused; 1.
FT	DOMAIN 20 31 POLY-ALA.	DR	InterPro: IPR01468; IGPS.
FT	DOMAIN 69 74 POLY-SER.	DR	InterPro: IPR001240; PRAT.
FT	DOMAIN 95 102 POLY-ALA.	DR	Pfam: PF00218; IGPS; 1.
FT	VARIANT 153 153 L -> F.	DR	Pfam: PF00697; PRAT; 1.
FT	CONFLICT 160 160 /FTI=VAR_010190.	DR	ProDom: P001511; IGPS; 1.
FT	SEQUENCE 1065 AA; 111996 MW; B84971AA7630A799 CRC34;	DR	ProDom: P006164; IGPS; 1.
DR	Query Match 22.3%; Score 57.5; DB 1; Length 1065; Best Local Similarity 25.9%; Pred. No. 28; Indels 11; Gaps 1; Matches 14; Conservative 9; Mismatches 20; ID 4; SEQ 4	KW	Tryptophan biosynthesis; Isomerase; Lyase; Multifunctional enzyme; Decarboxylase; Complete protease.
DR	Query Match 22.1%; Score 57; DB 1; Length 474; Best Local Similarity 25.6%; Pred. No. 14; ID 4; SEQ 4	KW	INDOLE-3-GLYCEROL PHOSPHATE SYNTHASE.
DR	Query Match 22.1%; Score 57; DB 1; Length 474; Best Local Similarity 25.6%; Pred. No. 14; ID 4; SEQ 4	FT	CONFLICT 88 89 N-(5'-PHOSPHORYL)ANTHRANILATE ISOMERASE.
DR	Query Match 22.1%; Score 57; DB 1; Length 474; Best Local Similarity 25.6%; Pred. No. 14; ID 4; SEQ 4	FT	CONFLICT 110 110 A -> G (IN REF. 1).
DR	Query Match 22.1%; Score 57; DB 1; Length 474; Best Local Similarity 25.6%; Pred. No. 14; ID 4; SEQ 4	FT	CONFLICT 130 131 H -> R (IN REF. 1).
DR	Query Match 22.1%; Score 57; DB 1; Length 474; Best Local Similarity 25.6%; Pred. No. 14; ID 4; SEQ 4	FT	CONFLICT 153 153 A -> D (IN REF. 1).
DR	Query Match 22.1%; Score 57; DB 1; Length 474; Best Local Similarity 25.6%; Pred. No. 14; ID 4; SEQ 4	FT	CONFLICT 302 302 L -> S (IN REF. 1).
DR	Query Match 22.1%; Score 57; DB 1; Length 474; Best Local Similarity 25.6%; Pred. No. 14; ID 4; SEQ 4	FT	CONFLICT 343 343 D -> G (IN REF. 1).
DR	Query Match 22.1%; Score 57; DB 1; Length 474; Best Local Similarity 25.6%; Pred. No. 14; ID 4; SEQ 4	FT	CONFLICT 381 383 MISSING (IN REF. 1).
DR	Query Match 22.1%; Score 57; DB 1; Length 474; Best Local Similarity 25.6%; Pred. No. 14; ID 4; SEQ 4	FT	CONFLICT 454 474 AGAKOGAGALIKIATSTFV -> GWGERCRRRAENFRDH.
DR	Query Match 22.1%; Score 57; DB 1; Length 474; Best Local Similarity 25.6%; Pred. No. 14; ID 4; SEQ 4	FT	INDIPLKVK (IN REF. 1).

Matches 10; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 13 VVEGLAGLEQLRARLEH 30
ID :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

Query Match 21.9%; Score 56.5; DB 1; Length 215;
Best Local Similarity 26.7%; Pred. No. 6.8;
Matches 16; Conservative 8; Mismatches 23; Indels 13; Gaps 1;

RESULT 5
GPMB_SALTY
ID GPMB_SALTY STANDARD; PRT; 215 AA.
AC Q8Z0T4;
DT 28-FEB-2003 (Rel. 41, last sequence update)
DT 10-OCT-2003 (Rel. 42, last annotation update)
DE Probable phosphoglycerate mutase gpmb (EC 5.4.2.1)
DE (phosphoglyceromutase) (PGAM).
OS GPNB OR STM4932 OR T4624.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
RN NCBITaxonID=602;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=TY2 / ATCC 700720;
RX MEDLINE=22331367; PubMed=12644504;
RA Deng W., Liou S.-R., Plunkett G., III, Mayhew G.F., Rose D.J.,
RA Burdell V., Kodoyianni V., Schwartz D.C., Blattner F.R.;
RT "Comparative genomics of *Salmonella enterica* serovar Ty2
and C₁B." J. Bacteriol. 185:2330-2337(2003).
CC -!- CATALYTIC ACTIVITY: 2-phospho-D-glycerate + 2,3-diphosphoglycerate
CC -!- 3-phospho-D-glycerate + 2,3-diphosphoglycerate.
CC -!- PATHWAY: Glycolysis.
CC -!- SIMILARITY: Belongs to the phosphoglycerate mutase family. Gpmb
CC subfamily.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
EMBL: AE018494; AA077055.1; -.
DR HAMAP; MF_01040; -.
DR InterPro; IPR001345; PG/BPGM_mutase.
PF Pfam; PF00300; PGAM; 1.
DR PROSITE; PS00175; PG_MUTASE; 1.
FT ACT_SITE 9 : Isomerase; Glycolysis; Complete proteome.
FT ACT_SITE 9 : TELE-PHOPHORISTIDINE INTERMEDIATE
FT ACT_SITE 9 : (BY SIMILARITY).
FT PROSITE; PS00175; PG_MUTASE; 1.
FT ACT_SITE 150 : Best Local Similarity 26.7%; Pred. No. 6.8;
FT ACT_SITE 150 : Matches 16; Conservative 8; Mismatches 23; Indels 13; Gaps 1;
FT ACT_SITE 150 : BY SIMILARITY.
FT SEQUENCE 215 AA; 23958 MW; ADFFDD335F50F809 CRC64;
FT ACT_SITE 58 : REQUIRED FOR BINDING CARBOYL GROUP OF
FT ACT_SITE 58 : PHOSPHOCREATES (BY SIMILARITY).
FT SEQUENCE 215 AA; 23958 MW; BC470395F0FF819 CRC64;

RESULT 6
GPMB_SALTY
ID GPMB_SALTY STANDARD; PRT; 215 AA.
AC Q8Z0T4;
DT 28-FEB-2003 (Rel. 41, last sequence update)
DT 10-OCT-2003 (Rel. 42, last annotation update)
DE Probable phosphoglycerate mutase gpmb (EC 5.4.2.1)
DE (phosphoglyceromutase) (PGAM).
OS GPNB OR STM4585.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
RN NCBITaxonID=602;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=LT2 / SGSC112 / ATCC 700720;
RX MEDLINE=21534348; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott J., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Florea L., Milner W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of *Salmonella enterica* serovar Typhimurium
LT2." Nature 413:852-856(2001).
CC -!- CATALYTIC ACTIVITY: 2-phospho-D-glycerate + 2,3-diphosphoglycerate.
CC -!- PATHWAY: Glycolysis.
CC -!- SIMILARITY: Belongs to the phosphoglycerate mutase family. Gpmb
CC subfamily.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
EMBL: AE00915; AAL23400.1; -.
DR STYGene; SG?????; gpmb.
DR HAMAP; MF_01040; -.
DR InterPro; IPR001345; PG/BPGM_mutase.
PF Pfam; PF00300; PGAM; 1.
DR PROSITE; PS00175; PG_MUTASE; 1.
FT ACT_SITE 9 : Isomerase; Glycolysis; Complete proteome.
FT ACT_SITE 9 : TELE-PHOPHORISTIDINE INTERMEDIATE
FT ACT_SITE 9 : (BY SIMILARITY).
FT PROSITE; PS00175; PG_MUTASE; 1.
FT ACT_SITE 150 : Best Local Similarity 26.6%; Pred. No. 6.8;
FT ACT_SITE 150 : Matches 16; Conservative 8; Mismatches 23; Indels 13; Gaps 1;
FT ACT_SITE 150 : BY SIMILARITY.
FT SEQUENCE 215 AA; 23958 MW; ADFFDD335F50F809 CRC64;
FT ACT_SITE 58 : REQUIRED FOR BINDING CARBOYL GROUP OF
FT ACT_SITE 58 : PHOSPHOCREATES (BY SIMILARITY).
FT SEQUENCE 215 AA; 23958 MW; BC470395F0FF819 CRC64;

RESULT 7

KVB3_HUMAN STANDARD; PRT; 404 AA.
 ID KVB3_HUMAN STANDARD; PRT; 404 AA.
 AC O3448; DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Voltage-gated potassium channel beta-3 subunit (K⁺ channel beta-3 subunit) (Kv-beta-3).
 DB KCNAB3 OR KCNAB3B.
 OS Homo sapiens (Human)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OX NEBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A., FUNCTION, AND TISSUE SPECIFICITY.
 RC TISSUE-Brain;
 RX MEDLINE:99074289; PubMed=9857044;
 RA Leicher T., Baehrung R., Ibbstrand D., Pongs O.;
 RT "Coexpression of the KCNAB gene product with Kv1.5 leads to a novel A-type potassium channel.";
 RL J. Biol. Chem. 273:35095-35101(1998).
 CC -!- FUNCTION: Accessory potassium channel protein which modulates the activity of the pore-forming alpha subunit. Alters the functional properties of Kv1.5.
 CC -!- SUBUNIT: Forms heteromultimeric complex with alpha subunits.
 CC -!- SUBCELLULAR LOCATION: Cyttoplasmic (Potential).
 CC -!- TISSUE SPECIFICITY: Brain-specific expression. Most prominent expression in cerebellum. Weaker signals detected in cortex, occipital lobe, frontal lobe and temporal lobe. Not detected in spinal cord, heart, lung, liver, kidney, pancreas, placenta and skeletal muscle.
 CC -!- DOMAIN: Alteration of functional properties of alpha subunit is mediated through N-terminal domain of beta subunit (Probable).
 CC -!- SIMILARITY: Belongs to the shaker potassium channel beta subunit family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or_send_an_email_to_license@isb-sib.ch).
 CC
 CC EMBL; AF016411; AAB92499.1; -.
 DR Gene: HGNC:6230; KCNAB3.
 DR GO: GO:0015459; R-type potassium channel regulator activity; TAS.
 DR GO: GO:0006813; R-type potassium ion transport; TAS.
 DR InterPro; IPR01395; Aldo/ket red.
 DR InterPro; IPR005402; KCNAB3 channel.
 DR InterPro; IPR03399; KCNAB channel.
 DR InterPro; IPR005983; KCNAB core.
 PRAM; P00248; aldo ket red; 1.
 PRINTS; PR01580; KCNAB CHANNEL.
 DR PRINTS; PR01577; KCNAB CHANNEL.
 DR PRODOM; PD000288; Aldo/ket red; 2.
 TIGRFAMS; TIGR01293; Kv_beta; 1.
 KW Ionic channel; Ion transport; Potassium transport;
 KW Voltage-gated channel.
 SQ SEQUENCE 404 AA; 43530 MW; 08265CC0729A1BA CRC64;
 Query Match 21.9%; Score 56.5; DB 1; Length 404;
 Best Local Similarity 31.1%; Pred. No. 13;
 Matches 14; Conservative 5; Mismatches 3; Indels 23; Gaps 2;

KVB3_HUMAN STANDARD; PRT; 415 AA.
 ID BFS2_BOVIN STANDARD; PRT; 415 AA.
 AC Q28177; DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Phakolin (beaded filament structural protein 2) (lens fiber cell beaded filament protein CP 49) (CP49) (49 kDa cytoskeletal protein) (CP 47) (CP47).
 DB BRSF2.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Lens;
 RX MEDLINE:94075412; PubMed=7504675;
 RA Merdes A., Gounari F., Georgatos S.D.;
 RT "The 47-kb lens-specific protein phakolin is a tailless intermediate filament protein and an assembly partner of filensin.";
 RL J. Cell Biol. 123:1507-1516(1993).
 RN [2]
 RP REVISORS.
 RA Merdes A.;
 RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
 CC -!- SUBUNIT: ASSOCIATES WITH BRSF1.
 CC -!- SUBCELLULAR LOCATION: Membrane- and cytoskeleton-associated.
 CC -!- TISSUE SPECIFICITY: Lens.
 CC -!- SIMILARITY: Belongs to the intermediate filament family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or_send_an_email_to_license@isb-sib.ch).
 CC
 CC EMBL; X75160; CMA53003.1; -.
 DR InterPro; IPR001664; IP: InterPro; IPR002957; Keratin_L.
 DR PFAM; PF00338; filament; 1.
 DR PRINTS; PR01248; TYPKERATIN.
 DR PROSITE; PS00245; IP: FALSE NEG.
 DR KW Intermediate filament; REPEAT; Membrane; Coiled coil; Cytoskeleton; Eye lens protein.
 DR DOMAIN 1 114 HEAD.
 FT DOMAIN 115 395 ROD.
 FT DOMAIN 396 415 TAIL.
 FT DOMAIN 115 144 COILED COIL (POTENTIAL).
 FT DOMAIN 199 248 COILED COIL (POTENTIAL).
 FT DOMAIN 295 395 COILED COIL (POTENTIAL).
 SQ SEQUENCE 415 AA; 45949 MW; F399685286966186 CRC64;
 Query Match 21.9%; Score 56.5; DB 1; Length 415;
 Best Local Similarity 41.7%; Pred. No. 14;
 Matches 15; Conservative 5; Mismatches 15; Indels 1; Gaps 1;
 QY 11 GHVVEGAGELFOLARLEHHPOGQRPMMGGCKL 46
 DB 364 GAVVSRLEAELEMRRBAAEOLQA-REHLLSHKCOL 398
 SNCX_HUMAN STANDARD; PRT; 1560 AA.
 ID SNCX_HUMAN STANDARD; PRT; 1560 AA.
 AC P41227; DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE SnCx protein (Kef169 protein).
 GN SNCX OR XE169.

CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; AL939126; CAB60167.1; -;
DR InterPro; IPR005164; Allantoicase.
DR Pfam; PF03561; Allantoicase.
KW Hypothetical protein; Hydrolase; Purine metabolism; Complete proteome.
SEQUENCE 376 AA; 41207 MW; 2247C0B2300C4A29 CRC64;
SQ

Query Match 20.9%; Score 54; DB 1; Length 376;
Best Local Similarity 44.4%; Pred. No. 26;
Matches 12; Conservative 1; Mismatches 12; Indels 2; Gaps 1;
QY 10 GGHVVEGLLAGLEO--LRARLEHHPOG 34
||| : |||
DB 171 GGHVANGFASVABQRFHRLRKQHPDG 197

Search completed: July 15, 2004, 20:34:28
Job time : 9.84706 secs

This Page Blank (uspto)

Copyright (c) 1993 - 2004 GenoScope version 5.1.6 CompuGen Ltd.

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Query	Match	28.5 %; Score 73.5; DB 16; Length 554;
	Best Local Similarity	35.4 %; Pred. No. 1.4; Mismatches 20; Indels 5; Gaps 1;
	Matches	17; Conservative
5	SEQUENCE	554 AA; 6466 MW; 030FF1804A7CDE80 CRC64;
TTGRRGHVET	GLAGELQRLARLHHHQGQRFPMMSGCKLG 47	
ALIGNMENTS		
RESULT 1		
Q9A9H1	Q9A9H1	PRELIMINARY
AC	Q9A9H1;	PRT: 554 AA.
DT	01-JUN-2001 (TREMBrel	17, Created)
DT	01-JUN-2001 (TREMBrel	17, Last sequence update)
DT	01-JUN-2003 (TREMBrel	24, Last annotation update)
DE	Hypothetical protein Cc1017.	
GN	CC1017.	
OS	Caulobacter crescentus.	
OC	Bacteria; Proteobacteria; Alpha proteobacteria; Caulobacterales;	
OC	Caulobacteraceae; Caulobacter.	
OX	NCBI_TAXID=155892;	
RN	[1]	
SEQUENCE FROM N.A.		
RC	STRAIN=ATCC 19089 / CB35;	
RX	MEDLINE:21173698; PubMed:11259647;	
RA	Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,	
RA	Petersen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,	
RA	Potenza I., Nelson W.C., Newton A., Stoeckli C., Phadke N.D., Ely B.,	
RA	DeBoy R.T., Dodson R.J., Durkin A.S., Swaminathan S., Gwinn M.L., Haft D.H.,	
RA	Kolonay J.F., Smit J.J., Craven M.B., Khouri H., Shetty J.J., Berry K.R., White O.,	
RA	Uttarbeck T., Tran K., Wolf A., Yamamethan J., Ermolaeva M., Fraser C.M.,	
RA	Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.,	
RT	"Complete genome sequence of Caulobacter crescentus ";	
RL	PROC. Natl. Acad. Sci. U.S.A. 98:4136-4141 (2001).	
DR	EMBL; AE005780; AAC23001.1; -;	
DR	PIR; E87375; E87375.	
DR	TIGR; CC1017; -;	
DR	Go:0008757; F1S-adenosylmethionine-dependent methyltransf. . . IEA.	
DR	InterPro; IPR00001; SAM bind.	
KW	Hypothetical protein; Complete proteome.	
SO	SEQUENCE	

RESULT 2	82BY9	PRELIMINARY;	PRT;	739 AA.
ID	Q82BY9			
AC	AC			
DT	01-JUN-2003 (TREMBLrel. 24, Created)			
DT	01-JUN-2003 (TREMBLrel. 24, Last sequence update)			
DE	putative regulatory protein.			
GN	SAV5565.			
OS	Streptomyces avermitilis.			
OC	Bacteria; Actinobacteria; Actinomycetidae; Actinomycetales;			
OC	Streptomycineae; Streptomycetaceae; Streptomyces.			
OX	NCBI_TAXID=33903;			
RN	SEQUENCE FROM N.A.			
RC	STRAIN=MA-680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;			
RX	MEDLINE=21477403; PubMed=11572988;			
RX	Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,			
RX	Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osonee T.,			
RX	Kikuchi H., Shiba T., Sakai Y., Hattori M.;			
RT	"Genome sequence of an industrial microorganism Streptomyces avermitilis; deducing the ability of producing secondary metabolites";			
RT	Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220 (2001).			
RN	[12]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=MA-680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;			
RX	MEDLINE=22608306; PubMed=12692562;			
RX	Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,			
RX	Sakai Y., Hattori M., Omura S.;			
RT	"Complete genome sequence and comparative analysis of the industrial microorganism Streptomyces avermitilis.;"			
RT	Nat. Biotechnol. 21:526-531 (2003).			
RT	EMLB; AP05043; BAC73277.1--;			
RT	EMBL; GO:000524; F: catalytic activity; IEA.			
RT	GO:0003824; F:catalytic activity; IEA.			
RT	InterPro; IPR03594; ATPbind ATPase.			
RT	InterPro; IPR01932; PP2C-like.			
DR	PFam; PF02518; HATPase c; 1.			
DR	SM00331; PP2C_SIG; 1.			
KW	Complete proteome.			
SQ	SEQUENCE 739 AA; 77478 MW; 638E3F93EC67A2B6 CRC64;			
Query Match	25.8%; Score 66.5; DB 16; Length 739;			
Best Local Similarity	32.6%; Pred. No. 14;			
Matches	14; Conservative 5; Mismatches 13; Indels 11; Gaps 1;			
Qy	1 MDEKTTGMRG-----GHVVEGLAGELQRARLEHHP 32			
Db	341 LEDEATGWRGPERGMAPGPRLLARVWHGSENRIEDLRLRHP 383			
RESULT 3	08KRD7	PRELIMINARY;	PRT;	189 AA.
ID	Q8KRD7			
AC	Q8KRD7;			
DT	01-OCT-2002 (TREMBLrel. 22, Created)			
DT	01-OCT-2002 (TREMBLrel. 22, Last sequence update)			
DE	Transcription repressor KifA-like protein.			
OS	Rhodococcus erythropolis.			
OC	Bacteria; Actinobacteria; Actinomycetidae; Actinomycetales;			
OC	Corynebacterineae; Nocardiaceae; Rhodococcus.			
OK	NCBI_TAXID=1833;			
RN	SEQUENCE FROM N.A.			
RP	STRAIN=MP50;			
RX	MEDLINE=22083458; PubMed=12089004;			
RX	Trot S., Burger S., Calaminus C., Stoltz A.;			
RESULT 4	Q9A365	PRELIMINARY;	PRT;	528 AA.
ID	Q9A365			
AC	Q9A365;			
DT	01-JUN-2001 (TREMBLrel. 17, Created)			
DT	01-JUN-2001 (TREMBLrel. 17, Last sequence update)			
DR	Hypothetical protein CC3341.			
GN	CC3341.			
OS	Caulobacter crescentus.			
OC	Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacterales;			
OC	Caulobacteraceae; Caulobacter.			
OX	NCBI_TAXID=155892;			
RN	[11]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=ATCC 19089 / CB15;			
RX	MEDLINE=21173598; PubMed=11259647;			
RX	Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,			
RA	Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,			
RA	Potocka I., Nelson W.C., Newcom A., Steppens C., Phadke N.D., Elly			
RA	DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,			
RA	Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K., Utterback T., Tran K., Wolf A., Yamathene J., Bimolaeva M., Witt			
RA	Salzberg S.L., Venter J.C., Shapiro L., Fraser L., Fraser M., White			
RT	"Complete genome sequence of Caulobacter crescentus.;"			
RT	Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141 (2001).			
RL	DR			
DR	EMBL; AE010598; AX25303.1; -.			
DR	PIR; C87663; C87663.			
DR	TIGR; CC3341; -.			
DR	GO; GO:0008168; F:methyltransferase activity; IEA.			
DR	InterPro; IPR01601; Methyltransf.			
KW	Hypothetical protein; Complete proteome.			
SQ	SEQUENCE 528 AA; 57050 MW; E18A1F4DA41785D5 CRC64;			
Query Match	25.4%; Score 65.5; DB 16; Length 528;			
Best Local Similarity	37.8%; Pred. No. 12;			
Matches	17; Conservative 5; Mismatches 18; Indels 5; Gaps 1;			
Qy	8 WRGGTGTW-----GLAGELQRARLEHHPQOREPMMGCGKLG 47			
Db	23 WSAGYVETGTYGCPPELNPVRAALPLIAGRFRAPRIETACELG 67			
RESULT 5	Q9RDU1	PRELIMINARY;	PRT;	792 AA.
ID	Q9RDU1			
AC	Q9RDU1;			
DT	01-MAY-000 (TREMBLrel. 13, Created)			
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)			
DT	01-OCT-2003 (TREMBLrel. 25, Last annotation update)			
DR	Hypothetical protein SCO2560.			
GN	SCO2560 OR SCCT7.27C.			
OS	Streptomyces coelicolor.			
OC	Bacteria; Actinobacteria; Actinomycetidae; Actinomycetales;			
OC	Streptomycineae; Streptomycetaceae; Streptomyces.			
OX	NCBI_TAXID=1902;			
RN	SEQUENCE FROM N.A.			

RC STRAIN-A3 (2);
 RA Oliver K., Harris D.;
 RL Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
 RN [2]
 RT SEQUENCE FROM N.A.
 RC STRAIN-A3 (2);
 RA Parkhill J., Barrell B.G., Rajandream M.A.;
 RL Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
 RP
 RC STRAIN-A3 (2);
 RA Redenbach M., Kieser H.M., Denapaita D., Bichner A., Cullum J.,
 RA Kinashi H., Hopwood D.A.;
 RT "A set of ordered cosmids and a detailed genetic and physical map for
 the 8 Mb Streptomyces coelicolor A3 (2) chromosome.";
 RT Mol. Microbiol. 21:77-96 (1996).
 RL [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN-A3 (2) / M145;
 RA BENTLEY S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
 RA Cronin N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
 RA Huang C.-H., Kieber T., Larke L., Murphy L., Oliver K., O'Neil S.,
 RA Robbinsch E., Rajandream M.A., Rutherford K., Rutter S.,
 RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
 RA Warren T., Wietzorek A., Woodward J., Barrell B.G., Parkhill J.,
 RA Hopwood D.A.;
 RT "Complete genome sequence of the model actinomycete Streptomyces
 coelicolor A3 (2).";
 RL Nature 417:141-147 (2002).
 DR EMBL; AL99113; CAB66238.1; -;
 DR GO; GO:000524; F:ATP binding; IEA.
 DR GO; GO:000824; F:catalytic activity; IEA.
 DR InterPro; IPR003594; ATPbind_ATPase.
 DR InterPro; IPR003018; GAF.
 DR Pfam; PF02518; HATPase_c; 1.
 DR SMART; SW00065; GAF; 1.
 DR SMART; SW00331; HATPase_c; 1.
 DR SMART; SW00331; PP2C_SIG; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 792 AA; 8140 MW; EFE7CC6D6852D3D CRC64;

Query Match 24.0%; Score 62; DB 16; Length 792;
 Matches 14; Conservative 3; Mismatches 12; Indels 6; Gaps 1;

Qy 8 WRGCHWVVBGLAGLQEQLRARLEHHPOGQRPPMSG 42
 Db 422 WHSGE-----GHLDELRALEKDPPRPDRPWRSG 450

RESULT 6

O9K87 PRELIMINARY; PRT; 429 AA.

ID O9K87 PRELIMINARY; PRT; 429 AA.

AC O9K87; DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE Maltose/maltodextrin transport system (Maltose/maltodextrin-binding
 protein).
 DE protein).
 GN B02926.
 OS *Bacillus halodurans*.
 OC Bacteria; Fimicutes; Bacillales; Bacillaceae; Bacillus.
 RN NCBI_TaxID=86665;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-C-125 / JCM 9153;
 RC MEDLINE=2052582; Published=11058122;
 RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,

RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
 RA Horikoshi K.;
 RT "Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and genomic sequence comparison with *Bacillus subtilis*.";
 RT Nucleic Acids Res. 28:4317-4331 (2000).
 RL EMBL; AP001517; BAB06645.1; -.
 DR PIR; FB4015; FB4015.
 DR HSSP; P02928; ITUD.
 DR GO; GO:005363; F:maltose transporter activity; IEA.
 DR GO; GO:005215; F:transporter activity; IEA.
 DR GO; GO:0015768; P:maltose transport; IEA.
 DR GO; GO:006810; P:transport; IEA.
 DR InterPro; IPR006060; Maltose_BP.
 DR InterPro; IPR000437; Prok_lipoprot_S.
 DR InterPro; IPR006609; SBP_Eac1.
 DR InterPro; IPR006061; SBP_dom1.
 DR Pfam; PF01547; SBP_bac_1.1.
 DR PRINTS; PRO0180; MALTOSBP.
 DR PROSITE; PS00013; PROKAR_LIPOPROTEIN_1.
 DR PROSITE; PS0107; SBP_BACTERIAL_1.
 KW Complete proteome.
 SQ SEQUENCE 429 AA; 46536 MW; 5B17521A0FA19297 CRC64;

Query Match 23.6%; Score 61; DB 16; Length 429;
 Best Local Similarity 36.8%; Pred. No. 35;
 Matches 14; Conservative 8; Mismatches 16; Indels 0; Gaps 0;
 Qy 10 GSHVVEGLAEQLRARLEHHPOGQRPPMSGGGKGL 47
 Db 122 GDIYLGCLAAEELTPHQEGYPEGAGDADNYEGAGL 159

RESULT 7

O33541 PRELIMINARY; PRT; 773 AA.

ID O33541 PRELIMINARY; PRT; 773 AA.

AC O33541; DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE Sensor protein RSSA.
 GN RSSA.
 OS endosymbiont of *Riftia pachyptila*.
 OC Bacteria; Proteobacteria; Gammaproteobacteria;
 OC sulfur-oxidizing symbionts.
 [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97438512; PubMed=9293000;
 RN Hughes D.S., Felbeck H., Stein J.L.;
 RT "A histidine protein kinase homolog from the endosymbiont of the
 RT hydrothermal vent tubeworm *Riftia pachyptila*.";
 RL Appl. Environ. Microbiol. 63:3494-3498 (1997).
 CC -1- SIMILARITY: TO OTHER PROKARYOTIC SENSORY TRANSDUCTION HISTIDINE
 CC KINASES.
 DR EMBL; U93704; AAB71131.1; -.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:000524; F:ATP binding; IEA.
 DR GO; GO:003677; F:DNA binding; IEA.
 DR GO; GO:0016301; F:kinase activity; IEA.
 DR GO; GO:0016740; F:transfase activity; IEA.
 DR GO; GO:000156; F:two-component response regulator activity; IEA.
 DR GO; GO:000155; F:two-component sensor molecule activity; IEA.
 DR GO; GO:000160; F:two-component signal transduction system (p. . .); IEA.
 DR InterPro; IPR003594; ATPbind_Atpase.
 DR InterPro; IPR005467; His_Kinase.
 DR InterPro; IPR003661; His_KinA_N.
 DR Pfam; PF02518; HATPase_C; 1.
 DR Pfam; PF00512; HisKA; 1.
 DR Pfam; PF00072; response_reg; 1.
 DR PRINTS; PRO0344; BCTRLSENSOR.

DR	InterPro; IPR007645; RNA_Pol_Rpb2-3.
DR	InterPro; IPR007120; RNA_Pol_Rpb2-6.
DR	InterPro; IPR007641; RNA_Pol_Rpb2-7.
DR	PFam; PF04583; RNA_Pol_Rpb2-1; 1.
DR	PFam; PF04561; RNA_Pol_Rpb2-2; 1.
DR	PFam; PF04565; RNA_Pol_Rpb2-3; 1.
DR	PFam; PF00562; RNA_Pol_Rpb2-6; 1.
DR	PFam; PF04560; RNA_Pol_Rpb2-7; 1.
DR	PROSITE; PS01166; RNA_Pol_BBTA; 1.
KW	Complete proteome.
SEQUENCE	1181 AA; 130439 MW; AEBF77ACC879EDBB CRC64;
Qy	4 KTTGWRGGHVGELAGELBQLRALLE-HHPOGOREPPM 40
Db	247 KALGWTTTEQILEER-GEYESMRATLEKQDHTQGDDALL 283
RESULT 10	
ID	O9RKG2
AC	OPRKG2
DT	01-MAY-2000 (TREMBLrel. 13, Created)
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT	01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE	putative isomerase.
GN	SC03475 OR SCE65_11C OR SC656_11C.
OS	Streptomyces coelicolor.
OC	bacteria; Actinobacteria; Actinomycetales;
OC	Streptomyces; Streptomycetaceae; Streptomyces.
[1]	NCBI_TaxID=1902;
RP	SEQUENCE FROM N.A.
RC	STRAIN-A3(2);
RA	Murphy L., Harris D.; to the EMBL/GenBank/DDBJ databases.
RL	Submitted (Nov-1999) to the EMBL/GenBank/DDBJ databases.
RN	[2]
RP	SEQUENCE FROM N.A.
RC	STRAIN-A3(2);
RA	Cerdeno A.M., Parkhill J., Barrell B.G., Rajandream M.A.;
RL	Submitted (Nov-1999) to the EMBL/GenBank/DDBJ databases.
RN	[3]
RP	SEQUENCE FROM N.A.
RC	STRAIN-A3(2);
RA	Redenbach M., Kieser H.M., Denapaité D., Eichner A., Cullum J.,
RA	Kinasho H., Hopwood D.A.;
RT	"A set of ordered cosmids and a detailed genetic and physical map for the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RT	the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL	Microbiol. 21:77-96(1996).
RN	[4]
RP	SEQUENCE FROM N.A.
RC	STRAIN-A3(2) / MI:45;
RA	Medline:2198641; PubMed:12000953;
RA	Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA	Thomson M.R., James K.D., Harris D.E., Quail M.A., Kieser H., Hails M.,
RA	Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA	Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA	Huang C.-H., Kleiser T., Larke L., Murphy L., Oliver K., O'Neil S.,
RA	Rabinowitzsch E., Rajandream M.A., Rutherford K., Rutter S.,
RA	Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA	Warren T., Wietzorek A., Woodward J., Barrell B.G., Parkhill J.,
RA	Hopwood D.A.;
RT	"Complete genome sequence of the model actinomycete Streptomyces coelicolor A3(2)." ;
RL	Nature 417:141-147 (2002).
DR	EMBL; AL939116; CAB61799.1; -.
DR	HSSP; P42206; IBQC.
DR	GO; GO:001683; F:isomerase activity; IEA.
DR	GO; GO:0008152; P:metabolism; IEA.
DR	
DR	InterPro; IPR001354; MR_MLE.
DR	PFam; PF01188; MR_MLE; 1.
DR	PFam; PF02746; MR_MLE_N; 1.
DR	PROSITE; PS00908; MR_MLE; 1.
DR	PROSITE; PS00905; MR_MLE; 2.
KW	Isomerase; Complete Proteome.
SEQUENCE	377 AA; 40887 MW; 8787DD4A30985F91 CRC64;
Qy	23.6%; Score 61; DB 16; Length 1181;
Best Local Similarity	39.5%; Pred. No. 1.1e+02;
Matches	15; Conservative 6; Mismatches 15; Indels 2; Gaps 2;
Db	23 DKGVIGN-GRAVVEQGARRVVERMIGDIAEVLAGRDPRRIEHHQGMRGAFYRGCAVL 79
RESULT 11	
ID	Q8D3W1
AC	Q8D3W1
DT	01-MAR-2003 (TREMBLrel. 23, Created)
DT	01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DB	ABC-type protease/lipase transport system, ATPase and permease component.
GN	VV21572.
OS	Vibrio vulnificus.
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales; Vibrionaceae; Vibrio.
OX	NCBI_TaxID=672;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN-CMC6;
RA	Rhee J.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H.,
RA	Choy H.E.;
RT	"Complete genome sequence of vibrio vulnificus CMCP6.;"
RL	Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.
DR	EMBL; AE016813; AA008435.1; -.
DR	GO; GO:016020; C:membrane; IEA.
DR	GO; GO:000524; F:ATP binding; IEA.
DR	GO; GO:0004040; F:ABC transporter acti. . .; IEA.
DR	GO; GO:000166; F:nucleotide binding; IEA.
DR	GO; GO:000233; F:peptidase activity; IEA.
DR	GO; GO:0006810; P:transport; IEA.
DR	InterPro; IPR003593; AAA_ATPase.
DR	InterPro; IPR004349; ABC transporter.
DR	PFAM; PF00005; ABC tran.; 1.
SMART; SM00382; AAA; 1.	
PROSITE; PS00211; ABC TRANSPORTER_1; 1.	
PROSITE; PS00833; ABC TRANSPORTER_2; 1.	
DR	Protease; Complete Proteome.
KW	SEQUENCE 420 AA; 45496 MW; 98A77D0FA3E3A51B CRC64;
Qy	Query Match 23.3%; Score 60; DB 16; Length 420;
Best Local Similarity	21.1%; Pred. No. 46;
Matches	16; Conservative 13; Mismatches 17; Indels 30; Gaps 3;
Db	137 MDEKTTGWRGGHVGELAGELBQLRALLE-----HHPOGOREP 38
Qy	39 MMSG-----GCKLG 47
Db	196 VLSGAOFSPPPGCAVG 211
RESULT 12	
ID	Q8DH85
AC	Q8DH85
DT	01-MAR-2003 (TREMBLrel. 23, Last sequence update)

DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)

DE Penicillin-binding protein.

GN TLR274.

OS Synechococcus elongatus (Thermosynechococcus elongatus).

OC Bacteria; Cyanobacteria; Chroococcales; Synechococcus.

RN [1] NCBI_TaxID:32046;

RP SQUENCE FROM N.A.

RX STRAIN=BP-1;

MEMLINE=2225144; Pubmed=12240834; GO; GO:0005215; F:transporter activity; IEA.

RN Nakamura Y., Kaneko T., Sato S., Ikeuchi M., Katoh H., Sasamoto S., Kiyokawa C., Kohara M., Matsunoto M., Matsuno A., Nakazaki N., Shimpo S., Sugimoto M., Takeuchi C., Yamada M., Tabata S.; "Complete genome structure of the thermophilic cyanobacterium Thermosynechococcus elongatus BP-1."; DNA Res. 9;123-130(2002).

RN EMBL; AP005376; BAC05626.1; -;

DR GO; GO:008658; P:penicillin binding; IEA.

DR GO; GO:009273; P:cell wall biosynthesis (sensu Bacteria); IEA.

DR InterPro; IPR005311; PBP dimer.

DR Pfam; PF03717; PBP dimer; I.

DR Kw Complete proteome.

SQ SQUENCE - 588 AA; 65048 MW; 9E24A8544A102719 CRC64;

Query Match 23.3%; Score 60; DB 16; Length 588; Best Local Similarity 57.1%; Pred. No. 67; Matches 12; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 12 HVVEGLAGLEBLQIARLLEHH 32

Db 442 HVIDGLYDENGQQLKRLEHP 462

RESULT 13

ID 08XAY8 PRELIMINARY; PRT; 342 AA.

AC 08XAY8; DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)

DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)

DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)

DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)

DE Putative transport system permease protein.

GN YDBY OR 22191 OR ECS2121.

OS Bacteriobacteri coli 0157:H7.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Bacteriobacteri.

OC NCBI_TaxID=83334;

OX NCBI_TaxID=83334;

RN [1] NCBI_TaxID=83334;

RP SQUENCE FROM N.A.

RC STRAIN=0157:H7 / EDL933 / ATCC 700927;

RX MEDLINE=21074935; Pubmed=11206551; Perna N.T., Plumbett G. III, Burland V., Mau B., Glasner J.D., Rose D. J., Mayhew G.F., Evans P.S., Gregor J., Karppatrick H.A., Rosai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L., Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamitis K., Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C., Welch R.A., Blattner F.R., "Genome sequence of enterohemorrhagic Escherichia coli 0157:H7.,"; Nature 409;529-533 (2001).

RN [2] SEQUENCE FROM N.A.

RC STRAIN=0157:H7 / RIMD 0509952; MEDLINE=21156231; Pubmed=11258796; RY Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T., Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T., Kubara S., Shiba T., Hattori M., Shinagawa H.; "Complete genome sequence of enterohemorrhagic Escherichia coli 0157:H7 and genomic comparison with a laboratory strain K-12.,"; DNA Res. 8;11-22(2001).

RN EMBL; AE005353; AAG56252.1; -.

DR EMBL; AP002557; BAB35544.1; -;

DR PIR; A90834; A90894;

DR PIR; H85723; H85723.

DR GO; GO:0016020; C:membrane; IEA.

DR GO; GO:0005215; F:transporter activity; IEA.

DR GO; GO:0006810; P:transport; IEA.

DR InterPro; IPR0011851; Bac_inmem_transp.

DR Pfam; PF02653; BPD_transp_2; I.

RN Kw Complete proteome.

SQ SQUENCE - 342 AA; 36365 MW; 6953A23C07C305AB CRC64;

Query Match 23.1%; Score 59.5; DB 16; Length 342; Best Local Similarity 35.0%; Pred. No. 42; Matches 14; Conservative 5; Mismatches 10; Indels 11; Gaps 1;

QY 8 WRGGHIVVEGLAGLEBLQIARLLEHHPOGQREPPMSGGCKLG 47

Db 135 WNGKWKIEGJPAELKLSA-----PLIUSVAIG 163

RESULT 14

ID 096MD1 PRELIMINARY; PRT; 139 AA.

AC 096MD1; DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)

DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)

DB Hypothetical protein FLJ32575.

OS Homo sapiens (Human).

OC Bukarvota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo. NCBI_TaxID=9606;

RN [1] NCBI_TaxID=9606;

RP SQUENCE FROM N.A.

RC TISSUE=Spleen; Kawakami B., Sugiyama A., Takemoto M., Sugiyama T., Irie R., Otsuki T., Sato H., Ota T., Wakamatsu A., Ishii S., Yamamoto J., Isomura Y., Kawai-Hio Y., Saito K., Nishikawa T., Sekine T., Kimura K., Yamashita H., Matsuo K., Nakamura Y., Kikuchi H., Kanda K., Matsunaga M., Murakawa K., Kanehori K., Takahashi-Fujii A., Oshima A., Suzuki Y., Sugano S., Nagahara K., Masuno Y., Nagai K., Isogai T.; RT "NEDO human cDNA sequencing project"; RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases. DR EMBL; AK057137; BAB11369.1; -.

RN Hypothetical protein.

SQ SQUENCE - 139 AA; 15014 MW; 3A59D86197B48F19 CRC64;

Query Match 22.9%; Score 59; DB 4; Length 139; Best Local Similarity 44.8%; Pred. No. 18; Matches 13; Conservative 3; Mismatches 11; Indels 2; Gaps 1;

QY 3 EKTTGGRGRGHVVEGLAGLEBLQIARLLEHH 31

Db 92 ERATGWRSLHVEGCGSQGQTSCH--HH 118

RESULT 15

ID 08L5B8 PRELIMINARY; PRT; 149 AA.

AC 08L5B8; DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)

DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)

DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)

DE P0671D01.7 protein (P0445H04.30 protein).

GN P0671D01.7 OR P0445H04.30 protein.

OS Oryza sativa (Japonica cultivar-group).

OC Bukarvota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.

RC STRAIN=cv. Nipponbare;

This Page Blank (uspto)

SQ	Sequence 47 AA:	Best Local Similarity 100.0%; Pred. No. 2.3e-28; Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Query Match	100.0%; Score 258; DB 5; Length 47;	
Best Local Similarity	100.0%; Pred. No. 2.3e-28;	
Matches	47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1 MDEKTTGWRGHHVVEGLAGLEQQLRARLHHHQGOREPPMMGGCKLG 47	
Db	1 MDEKTTGWRGHHVVEGLAGLEQQLRARLHHHQGOREPPMMGGCKLG 47	
RESULT 2		
ABU57601		
ID ABU57601 standard; peptide; 47 AA.		
AC		
XX ABU57601;		
XK		
DT 09-APR-2003 (first entry)		
DE Synthetic peptide SB19 C-terminal deletion mutant, C4.		
XX Streptavidin-binding peptide; mRNA display; peptide library;		
KW fusion protein; mutant; mutein; SB19.		
XX OS Synthetic.		
XX DE		
XX Synthetic peptide SB19 truncation mutant #6.		
XX DE Streptavidin-binding peptide; mutant; mutein.		
XX KW Streptavidin-binding peptide; mutant; mutein.		
XX OS Escherichia coli.		
XX PR WO200238580-A1.		
XX BN		
XX PA		
XX PD 16-MAY-2002.		
XX PR 31-OCT-2000; 2000WO-US041717.		
XX PR 31-OCT-2000; 2000WO-US041717.		
XX DR WO200238580-A1.		
XX PA (GEHO) GEN HOSPITAL CORP.		
XX PR Szostak JW, Wilson DS, Keefe AD;		
XX PT		
XX DR WPI; 2002-500155/53.		
XX PS Novel peptide with high affinity for streptavidin, is expressed as part of fusion protein to facilitate detection, quantitation and purification of desired protein.		
XX PS WPI; 2003-182639/18.		
XX PS Claim 9; Fig 5; 56pp; English.		
XX CC The invention relates to a peptide which binds streptavidin with a dissociation constant less than 10 micro M, where the peptide is not disulphide bonded or cyclised, or with a dissociation constant less than 10 micro M, where the amino acid sequence of the peptide does not contain an HPO, HPN or HQP motif. The peptides are generated by the mRNA display method to create a library of peptides. Also include are a fusion protein encoding the peptide, a vector comprising the nucleic acid, a fusion protein (and its encoding nucleic acid/vector) comprising a protein of interest covalently linked to the streptavidin binding peptide and detecting the presence of the fusion protein in a sample using streptavidin or streptavidin-containing compounds. The fusion protein (which comprises the desired protein of interest covalently linked to the streptavidin binding peptide) is useful for purifying a protein of interest from a sample, which involves expressing the fusion protein in the sample, contacting the sample with streptavidin under conditions that allow complex formation between the fusion protein and the streptavidin, isolating the complex and recovering the fusion protein, and thus purifying the desired protein from the sample. The invention describes a method for producing a streptavidin-binding fusion protein. This Sequence represents a truncation mutant of the streptavidin-binding peptide SB-19 isolated from Escherichia coli using a method described in the specification.		
XX SQ Sequence 47 AA:		
Query Match 94.2%; Score 243; DB 5; Length 47;		
Best Local Similarity 95.7%; Pred. No. 2.8e-26;		
Matches 45; Conservative 0; Mismatches 2; Indels 0; Gaps 0;		
Qy 1 MDEKTTGWRGHHVVEGLAGLEQQLRARLHHHQGOREPPMMGGCKLG 47		
Db 1 MDEKTTGWRGHHVVEGLAGLEQQLRARLHHHQGOREPPMMGGCKLG 47		
RESULT 4		
ABU57602		
ID ABU57602 standard; peptide; 47 AA.		
XX AC ABU57602;		
XX SQ Sequence 47 AA:		
Query Match 100.0%; Score 258; DB 6; Length 47;		
Best Local Similarity 100.0%; Pred. No. 2.3e-28;		
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
Qy 1 MDEKTTGWRGHHVVEGLAGLEQQLRARLHHHQGOREPPMMGGCKLG 47		
Db 1 MDEKTTGWRGHHVVEGLAGLEQQLRARLHHHQGOREPPMMGGCKLG 47		
RESULT 3		
ABG67080		
ID ABG67080 standard; peptide; 47 AA.		
AC ABG67080;		
XX DT 24-SEP-2002 (first entry)		
XX DE Streptavidin-binding peptide SB19 truncation mutant #6.		
XX KW Streptavidin-binding peptide; mutant; mutein.		
XX OS Escherichia coli.		
XX PR		
XX PA (GEHO) GEN HOSPITAL CORP.		
XX PR Szostak JW, Wilson DS, Keefe AD;		
XX PT		
XX DR WPI; 2002-500155/53.		
XX PS Novel peptide with high affinity for streptavidin, is expressed as part of fusion protein to facilitate detection, quantitation and purification of desired protein.		
XX PS WPI; 2003-182639/18.		
XX PS Claim 9; Fig 5; 56pp; English.		
CC The invention describes a peptide (I) which binds streptavidin with a dissociation constant less than 10 mM or 23 nM, where the amino acid sequence of (I) does not contain an HPO, HPN or HQP motif, is not disulphide bonded or cyclised. A fusion protein comprising a protein of interest covalently linked to (I) is useful for purifying a desired protein from a sample by contacting the sample with streptavidin under conditions that allow complex formation between the fusion protein and the streptavidin, isolating the complex and recovering the fusion protein, and thus purifying the desired protein from the sample. The invention describes a method for producing a streptavidin-binding fusion protein. This Sequence represents a truncation mutant of the streptavidin-binding peptide SB-19 isolated from Escherichia coli using a method described in the specification.		
CC SQ Sequence 47 AA:		
Query Match 94.2%; Score 243; DB 5; Length 47;		
Best Local Similarity 95.7%; Pred. No. 2.8e-26;		
Matches 45; Conservative 0; Mismatches 2; Indels 0; Gaps 0;		
Qy 1 MDEKTTGWRGHHVVEGLAGLEQQLRARLHHHQGOREPPMMGGCKLG 47		
Db 1 MDEKTTGWRGHHVVEGLAGLEQQLRARLHHHQGOREPPMMGGCKLG 47		

DT	09-APR-2003	(first entry)	XX	Streptavidin-binding peptide; mutant; mutein.
XX			KW	
DE	Streptavidin-binding synthetic peptide SB19 point mutant M1.		XX	
XX			OS	Escherichia coli.
KW			XX	
	Streptavidin-binding peptide; mRNA display; peptide library;		PN	
KW	fusion protein; mutant; mutein; SB19.		XX	
XX			PD	W0200238580-A1.
OS			XX	
Synthetic.			16-MAY-2002.	
XX			XX	
PN	US2002155578-A1.		PF	31-OCT-2000; 2000WO-US041717.
XX			XX	
PD	24-OCT-2002.		PR	31-OCT-2000; 2000WO-US041717.
XX			XX	
PF	31-OCT-2001; 2001US-00004381.		PA	(GEHO) GEN HOSPITAL CORP.
XX			PI	Szostak JW, Wilson DS, Keefe AD;
PR	31-OCT-2000; 2000US-0244541P.		XX	
XX			DR	WPI; 2002-500155/53.
PA	(WILS/ WILSON D S.		XX	
PA	(KEEF/ KEFE A D.		PA	Novel peptide with high affinity for streptavidin, is expressed as part
XX			PT	of fusion protein to facilitate detection, quantitation and purification
PT	Szostak JW, Wilson DS, Keefe AD;		PT	of desired protein.
XX			XX	
DR	WPI; 2003-182639/18.		RS	Claim 9; Fig 5; 56pp; English.
XX			XX	
PT	Novel peptide which binds streptavidin with high affinity useful as		CC	The invention describes a peptide (I) which binds streptavidin with a
PT	affinity tags for purifying fusion proteins containing proteins of		CC	dissociation constant less than 10 micro M, where the peptide is not
PT	interest.		CC	sequence of (I) does not contain an HPQ, HPM, HPN or HOP motif, is not
XX			CC	disulfide bonded or cyclised. A fusion protein comprising a protein of
PS	Claim 14; Fig 5; 36pp; English.		CC	interest covalently linked to (I) is useful for purifying a desired
XX			CC	protein from a sample by contacting the sample with streptavidin under
CC	The invention relates to a peptide which binds streptavidin with a		CC	conditions that allow complex formation between the fusion protein and
CC	dissociation constant less than 10 micro M, where the peptide is not		CC	the streptavidin, isolating the complex and recovering the fusion
CC	disulphide bonded or cyclised, or with a dissociation constant less than		CC	protein, and thus purifying the desired protein from the sample. The
CC	10 micro M, where the amino acid sequence of the peptide does not contain		CC	invention describes a method for producing a streptavidin-binding fusion
CC	an HPQ, HPM, HPN or HOP motif. The peptides are generated by the mRNA		CC	protein. This sequence represents a truncation mutant of the streptavidin
CC	display method to create a library of peptides. Also include are a		CC	-binding peptide SB-19 isolated from Escherichia coli using a method
CC	vector comprising the nucleic acid, a nucleic acid/vector comprising a		CC	described in the specification
CC	protein of interest covalently linked to the streptavidin binding peptide		XX	
CC	and detecting the presence of the fusion protein in a sample using		SQ	Sequence 61 AA;
CC	streptavidin or streptavidin-containing compounds. The fusion protein		Query Match	93.4%; Score 241; DB 5; Length 61;
CC	(which comprises the desired protein of interest covalently linked to the		Best Local Similarity	77.0%; Pred. No. 7.3e-26; Mismatches 0; Indels 14; Gaps 1;
CC	streptavidin binding peptide) is useful for purifying a protein of		Matches	Conservative 47; Conservative 0; Mismatches 0; Indels 14; Gaps 1;
CC	interest from a sample, which involves expressing the fusion protein in		QY	1 MDEKTTGWRGGHVVVEGLAGELFOLRARLHHPOCGOREPMNSGGCKLG 46
CC	the sample, contacting the sample with streptavidin under conditions that		Db	1 MDEKTTGWRGGHVVVEGLAGELFOLRARLHHPOCGOREPMNSGGCKLG 60
CC	allow complex formation between the fusion protein and the streptavidin,		QY	47 G 47
CC	isolating the complex and recovering the fusion protein, thereby		Db	61 G 61
CC	purifying the protein of interest from the sample. The present sequence			
CC	is a mutated version of SB19, a streptavidin binding peptide of the			
CC	invention. The mutants comprise 3 N-terminal deletions, 3 C-terminal			
CC	deletions and a truncated N-terminal peptide with point mutations			
XX				
SQ	Sequence 47 AA;			
RESULT 5	Query Match	94.2%; Score 243; DB 6; Length 47;		RESULT 6
ABG67078	Best Local Similarity	95.7%; Pred. No. 2.8e-26;	ABU57600	
ID	Matches	Length 45; Conservative 0; Mismatches 2; Indels 0; Gaps 0;	ID	ABU57600 standard; peptide; 61 AA.
AC			XX	
XX			AC	ABU57600;
Db	QY	1 MDEKTTGWRGGHVVVEGLAGELFOLRARLHHPOCGOREPMNSGGCKLG 47	XX	
Db	1 MDEKTTGWRGGHVVVEGLAGELFOLRARLHHPOCGOREPMNSGGCKLG 47		DT	09-APR-2003 (first entry)
			XX	
			DE	Synthetic peptide SB19 C-terminal deletion mutant, C3.
			XX	
			KW	Streptavidin-binding peptide; mRNA display; peptide library;
			XX	
			OS	Streptavidin; mutant; mutein; SB19.
			XX	Synthetic.
			PN	US2002155578-A1.
			XX	
			PD	24-OCT-2002.
			XX	
			DE	Streptavidin-binding peptide SB19 truncation mutant #4.

31-OCT-2001; 2001US-00004381. PR XX
 31-OCT-2000; 2000US-0244541P. PR XX
 (S20S/) SZOSTAK J W. XX
 (WILS/) WILSON D S. XX
 (KEEF/) KEEFE A D. XX
 PI Szostak JW, Wilson DS, Keefe AD; DR XX
 XX DR WPI; 2003-182639/18. XX
 PT Novel peptide which binds streptavidin with high affinity useful as PT affinity tags for purifying fusion proteins containing proteins of PT interest. XX
 PS Claim 14; Fig 5; 36pp; English. XX
 CC The invention relates to a peptide which binds streptavidin with a CC dissociation constant less than 10 micro M, where the peptide is not CC disulfide bonded or cyclised, or with a dissociation constant less than CC 10 micro M, where the amino acid sequence of the peptide does not contain CC an HPQ, HPM, HPN or HQP motif. The peptides are generated by the mRNA CC display method to create a library of peptides. Also include are a CC nucleic acid encoding the peptide, a vector comprising the nucleic acid, CC a fusion protein (and its encoding nucleic acid/vector) comprising a CC protein of interest covalently linked to the streptavidin binding peptide CC and detecting the presence of the fusion protein in a sample using CC streptavidin or streptavidin-containing compounds. The fusion protein CC (which comprises the desired protein of interest covalently linked to the CC streptavidin binding peptide) is useful for purifying a protein of CC interest from a sample, which involves expressing the fusion protein in CC the sample, contacting the sample with streptavidin under conditions that CC allow complex formation between the fusion protein and the streptavidin, CC isolating the complex and recovering the fusion protein, thereby CC purifying the protein of interest from the sample. The present sequence CC is a mutated version of SB19, a streptavidin binding peptide of the CC invention. The mutants comprise 3 N-terminal deletions, 3 C-terminal CC deletions and a truncated N-terminal peptide with point mutations CC
 XX Sequence 61 AA;

Query Match 93.4%; Score 241; DB 6; Length 61;
 Matches 47; Conservative 0; Mismatches 0; Indels 14; Gaps 1;

QY 1 MDEKTTGWRGHHVVEGLAGELSQIRARLEHHPQGOREP-----MMSGCKL 46
 Db 1 MDEKTTGWRGHHVVEGLAGELSQIRARLEHHPQGOREP-----MMSGCKL 60

QY 47 G 47
 Db 61 G 61

RESULT 7
 ABG67077 ABG67077 standard; peptide; 75 AA.
 XX
 AC ABG67077;
 XX
 DT 24-SEP-2002 (first entry)
 DE Streptavidin-binding peptide SB19 truncation mutant #3.
 XX
 KW Streptavidin-binding peptide; mutant; mutein.
 OS Escherichia coli.
 XX
 PN WO200238580-A1..
 PD 16-MAY-2002.

Query Match 90.7%; Score 234; DB 5; Length 75;
 Matches 47; Conservative 0; Mismatches 0; Indels 28; Gaps 1;
 QY 1 MDEKTTGWRGHHVVEGLAGELSQIRARLEHHPQGOREP-----MMSGCKL 38
 Db 1 MDEKTTGWRGHHVVEGLAGELSQIRARLEHHPQGOREP-----MMSGCKL 75

QY 39 -----MMSGCKL 47
 Db 61 LIIDPVEAMMSGCKL 75

RESULT 8
 ABU57599 ABU57599 standard; peptide; 75 AA.
 ID ABU57599
 XX
 AC ABU57599;
 XX
 DT 09-APR-2003 (first entry)
 DE Synthetic Peptide SB19 C-terminal deletion mutant, c2.
 XX
 KW Streptavidin-binding peptide; mRNA display; peptide library;
 KW fusion protein; mutant; mutein; SB19.
 XX
 OS Synthetic.
 XX
 PN US200215578-A1.
 XX
 DD 24-OCT-2002.
 XX
 PR 31-OCT-2001; 2001US-00004381.
 XX
 PR 31-OCT-2000; 2000US-0244541P.
 XX
 PR (S20S/) SZOSTAK J W.
 PA (WILS/) WILSON D S.
 PA (KEEF/) KEEFE A D.
 XX
 PI Szostak JW, Wilson DS, Keefe AD;

31-OCT-2000; 2000WO-US041717. PR XX
 (GEHO) GEN HOSPITAL CORP. XX
 PI Szostak JW, Wilson DS, Keefe AD; DR XX
 XX DR WPI; 2002-500155/53. XX
 PT Novel peptide with high affinity for streptavidin, is expressed as part PT of fusion protein to facilitate detection, quantitation and purification PT of desired protein. XX
 PS Claim 9; Fig 5; 56pp; English. XX
 CC The invention describes a peptide (I) which binds streptavidin with a CC dissociation constant less than 10 μ M or 23 nM, where the amino acid CC sequence of (I) does not contain an HPQ, HPM, HPN or HQP motif, is not CC disulfide bonded or cyclised. A fusion protein comprising a protein of CC interest covalently linked to (I) is useful for purifying a desired protein from a sample by contacting the sample with Streptavidin under CC conditions that allow complex formation between the fusion protein and CC the streptavidin, isolating the complex and recovering the fusion CC protein, and thus purifying the desired protein from the sample. The CC invention describes a method for producing a streptavidin-binding fusion protein. This sequence represents a truncation mutant of the streptavidin CC -binding peptide SB-19 isolated from Escherichia coli using a method CC described in the specification. XX
 Sequence 75 AA;

Query Match 90.7%; Score 234; DB 5; Length 75;
 Matches 47; Conservative 0; Mismatches 0; Indels 28; Gaps 1;
 QY 1 MDEKTTGWRGHHVVEGLAGELSQIRARLEHHPQGOREP-----MMSGCKL 38
 Db 1 MDEKTTGWRGHHVVEGLAGELSQIRARLEHHPQGOREP-----MMSGCKL 75

Query Match 90.7%; Score 234; DB 5; Length 75;
 Matches 47; Conservative 0; Mismatches 0; Indels 28; Gaps 1;
 QY 1 MDEKTTGWRGHHVVEGLAGELSQIRARLEHHPQGOREP-----MMSGCKL 38
 Db 1 MDEKTTGWRGHHVVEGLAGELSQIRARLEHHPQGOREP-----MMSGCKL 75

RESULT 8
 ABU57599 ABU57599 standard; peptide; 75 AA.
 ID ABU57599
 XX
 AC ABU57599;
 XX
 DT 09-APR-2003 (first entry)
 DE Synthetic Peptide SB19 C-terminal deletion mutant, c2.
 XX
 KW Streptavidin-binding peptide; mRNA display; peptide library;
 KW fusion protein; mutant; mutein; SB19.
 XX
 OS Synthetic.
 XX
 PN US200215578-A1.
 XX
 DD 24-OCT-2002.
 XX
 PR 31-OCT-2001; 2001US-00004381.
 XX
 PR 31-OCT-2000; 2000US-0244541P.
 XX
 PR (S20S/) SZOSTAK J W.
 PA (WILS/) WILSON D S.
 PA (KEEF/) KEEFE A D.
 XX
 PI Szostak JW, Wilson DS, Keefe AD;

XX
DR WPI; 2003-182639/18.

XX
PT Novel peptide which binds streptavidin with high affinity useful as
PT affinity tags for purifying fusion proteins containing proteins of
PT interest.

XX
PS Claim 14; Fig 5; 36pp; English.

CC The invention relates to a peptide which binds streptavidin with a
CC dissociation constant less than 10 micro M, where the peptide is not
CC disulphide bonded or cyclised, or with a dissociation constant less than
CC 10 micro M, where the amino acid sequence of the peptide does not contain
CC an HPQ, HPM, HPN or HOP motif. The peptides are generated by the mRNA
CC display method to create a library of peptides. Also include are a
CC nucleic acid encoding the peptide, a vector comprising the nucleic acid,
CC a fusion protein (and its encoding nucleic acid/vector) comprising a
CC protein of interest covalently linked to the streptavidin binding peptide
CC and detecting the presence of the fusion protein in a sample using
CC streptavidin or streptavidin-containing compounds. The fusion protein
CC (which comprises the desired protein of interest covalently linked to the
CC streptavidin binding peptide) is useful for purifying a protein of
CC interest from a sample, which involves expressing the fusion protein in
CC the sample contacting the sample with streptavidin under conditions that
CC allow complex formation between the fusion protein and the streptavidin,
CC isolating the complex and recovering the fusion protein, thereby
CC purifying the protein of interest from the sample. The present sequence
CC is a mutated version of SB19, a streptavidin binding peptide of the
CC invention. The mutants comprise 3 N-terminal deletions, 3 C-terminal
CC deletions and a truncated N-terminal peptide with point mutations
XX SQ Sequence 75 AA;

Query Match 90.7%; Score 234; DB 6; Length 75;
Best Local Similarity 62.7%; Pred. No. 8.7e+25; Mismatches 0; Indels 0; Gaps 0;
Matches 47; Conservative 0; Mismatches 0; Indels 28; Gaps 1;

QY 1 MDEKTTGWRGGHVEGLAGELQARLLEHHPQGQREP----- 38
Db 1 MDEKTTGWRGGHVEGLAGELQARLLEHHPQGQREP----- 60

QY 39 -----MSGGCKLG 47
Db 61 LLDPPVEMMSGCKLG 75

QY 39 -----MSGGCKLG 47
Db 61 LLDPPVEMMSGCKLG 89

RESULT 9
ABC67076
ID ABC67076 standard; peptide; 89 AA.

AC ABC67076;
XX

DT 24-SEP-2002 (first entry)
XX
DE Streptavidin-binding peptide SB19 truncation mutant #2.
KW Streptavidin-binding peptide; mRNA display; peptide library;
XX
OS Synthetic.
XX
PN US2002155578-A1.

DD 24-OCT-2002.
XX
PF 31-OCT-2001; 2001US-00004381.

XX
PR 31-OCT-2000; 2000US-0244541P.

XX
OS Escherichia coli.
XX
PN WO200238580-A1.

DD 31-OCT-2000; 2000WO-US041717.
XX
PR 16-MAY-2002.

PP 31-OCT-2000; 2000WO-US041717.
XX
PR 31-OCT-2000; 2000WO-US041717.
XX
PR 31-OCT-2000; 2000WO-US041717.

XX
PR 31-OCT-2000; 2000WO-US041717.

XX
PA (GEHO) GEN HOSPITAL CORP.

XX
PI Szostak JW, Wilson DS, Keefe AD;

XX
DR WPI; 2002-500155/53.

XX
PT Novel peptide with high affinity for streptavidin, is expressed as part
PT of fusion protein to facilitate detection, quantitation and purification
PT of desired protein.

XX
PS Claim 9; Fig 5; 56pp; English.

CC The invention describes a peptide (I) which binds streptavidin with a
CC dissociation constant less than 10 mM or 23 nM, where the amino acid
CC sequence of (I) does not contain an HPQ, HPM, HPN or HOP motif, is not
CC disulphide bonded or cyclised. A fusion protein comprising a protein of
CC interest covalently linked to (I) is useful for purifying a desired
CC protein from a sample by contacting the sample with streptavidin under
CC conditions that allow complex formation between the fusion protein and
CC the streptavidin, isolating the complex and recovering the fusion
CC protein, and thus purifying the desired protein from the sample. The
CC invention describes a method for producing a streptavidin-binding fusion
CC protein. This sequence represents a truncation mutant of the streptavidin
CC -binding peptide SB19 isolated from Escherichia coli using a method
CC described in the specification

XX SQ Sequence 89 AA;

Query Match 88.0%; Score 227; DB 5; length 89;
Best Local Similarity 52.0%; Pred. No. 1e-23; Mismatches 0; Indels 42; Gaps 1;
Matches 47; Conservative 0; Mismatches 0; Indels 42; Gaps 1;

QY 1 MDEKTTGWRGGHVEGLAGELQARLLEHHPQGQREP----- 38
Db 1 MDEKTTGWRGGHVEGLAGELQARLLEHHPQGQREP----- 60

QY 39 -----MSGGCKLG 47
Db 61 LLDPPVEMMSGCKLG 89

RESULT 10
ABU57598
ID ABU57598 standard; peptide; 89 AA.
XX
AC ABU57598;
XX
DT 09-APR-2003 (first entry)
XX
DE Synthetic peptide SB19 C-terminal deletion mutant, Cl.

XX
Streptavidin-binding peptide; mRNA display; peptide library;

XX
KW Streptavidin-binding peptide; mutant; mutein; SB19.
XX
OS Synthetic.

XX
PN US2002155578-A1.

DD 24-OCT-2002.
XX
PF 31-OCT-2001; 2001US-00004381.

XX
PR 31-OCT-2000; 2000US-0244541P.

XX
OS Escherichia coli.
XX
PN (SZOS) SZOSTAK J W.
PA (WILS) WILSON D S.
PA (KEEF) KEEFE A D.

XX
PI Szostak JW, Wilson DS, Keefe AD;

XX
DR WPI; 2003-182639/18.

XX
PT Novel peptide which binds streptavidin with high affinity useful as
PT affinity tags for purifying fusion proteins containing proteins of
PT interest.

XX
PS Claim 14; Fig 5; 36pp; English.

CC The invention relates to a peptide which binds streptavidin with a
 CC dissociation constant less than 10 micro M, where the peptide is not
 CC disulphide bonded or cyclised, or with a dissociation constant less than
 CC 10 micro M, where the amino acid sequence of the peptide does not contain
 CC an HPO, HPM, HPN or HOP motif. The peptides are generated by the mRNA
 CC streptavidin method to create a library of peptides. Also include are a
 CC nucleic acid encoding the peptide, a vector comprising the nucleic acid,
 CC a fusion protein (and its encoding nucleic acid/vector) comprising a
 CC protein of interest covalently linked to the streptavidin binding peptide
 CC and detecting the presence of the fusion protein in a sample using
 CC streptavidin or streptavidin-containing compounds. The fusion protein
 CC (which comprises the desired protein of interest covalently linked to the
 CC streptavidin binding peptide) is useful for purifying a protein of
 CC interest from a sample, which involves expressing the fusion protein in
 CC the sample, contacting the sample with streptavidin under conditions that
 CC allow complex formation between the fusion protein and the streptavidin,
 CC isolating the complex and recovering the fusion protein, thereby
 CC purifying the protein of interest from the sample. The present sequence
 CC is a mutated version of SB19, a streptavidin binding peptide of the
 CC invention. The mutants comprise 3 N-terminal deletions, 3 C-terminal
 CC deletions and a truncated N-terminal peptide with point mutations
 XX Sequence 89 AA;

Query Match 89.0%; Score 227; DB 6; Length 89;
 Best Local Similarity 52.8%; Pred. No. 1e-23; 0; Mismatches 0; Indels 42; Gaps 1;
 Matches 47;

Qy 1 MDEKTTGWRGGHVGELAGLEQLEQARLLEHHPQGOREP----- 38
 Db 1 MDEKTTGWRGGHVGELAGLEQLEQARLLEHHPQGOREP----- 60

Qy 39 ----- MMSGCKLG 47
 Db 61 LLDPVVKLTDMWFKKFKNVSKDKRMTPYLEMYDWSGGCKLG 89

RESULT 11
 ABG67075
 ID ABG67075 standard; peptide; 101 AA.
 XX
 AC ABG67075;
 XX
 DT 24-SEP-2002 (first entry)

DE Streptavidin-binding peptide SB19 truncation mutant #1.
 KW Streptavidin-binding peptide; mutant; mutein.
 KW Escherichia coli.
 OS Synthetic.
 XX WO200238580-A1.
 PD 16-MAY-2002.
 PR 31-OCT-2000; 2000WO-US041717.

PR 31-OCT-2000; 2000WO-US041717.

PR (GEHO) GEN HOSPITAL CORP.
 PA Szostak JW, Wilson DS, Keefe AD;
 XX DR WPI; 2002-50015/53.
 XX Novel peptide with high affinity for streptavidin, is expressed as part
 PT of fusion protein to facilitate detection, quantitation and purification
 PT of desired protein.
 XX
 PS Claim 9; Fig 5; 56pp; English.
 XX
 CC The invention describes a peptide (I) which binds streptavidin with a
 CC dissociation constant less than 10 micro M, where the peptide is not
 CC disulphide bonded or cyclised, or with a dissociation constant less than
 CC 10 micro M, where the amino acid sequence of the peptide does not contain
 CC an HPO, HPM, HPN or HOP motif. The peptides are generated by the mRNA
 CC display method to create a library of peptides. Also include are a
 CC nucleic acid encoding the peptide, a vector comprising the nucleic acid,
 CC a fusion protein (and its encoding nucleic acid/vector) comprising a
 CC protein of interest covalently linked to the streptavidin binding peptide
 CC and detecting the presence of the fusion protein in a sample using
 CC streptavidin or streptavidin-containing compounds. The fusion protein
 CC (which comprises the desired protein of interest covalently linked to the
 CC streptavidin binding peptide) is useful for purifying a protein of
 CC interest from a sample, which involves expressing the fusion protein in
 CC the sample, contacting the sample with streptavidin under conditions that
 CC allow complex formation between the fusion protein and the streptavidin,
 CC isolating the complex and recovering the fusion protein, thereby
 CC purifying the protein of interest from the sample. The present sequence
 CC is a mutated version of SB19, a streptavidin binding peptide of the
 CC invention. The mutants comprise 3 N-terminal deletions, 3 C-terminal
 CC deletions and a truncated N-terminal peptide with point mutations
 XX Sequence 89 AA;

CC dissociation constant less than 10 μ M or 23 nM, where the amino acid
 CC sequence of (I) does not contain an HPO, HPM, HPN or HOP motif, is not
 CC disulphide bonded or cyclised. A fusion protein comprising a protein of
 CC interest covalently linked to (I) is useful for purifying a desired
 protein from a sample by contacting the sample with streptavidin under
 CC conditions that allow complex formation between the fusion protein and
 the streptavidin, isolating the complex and recovering the fusion
 protein, and thus purifying the desired protein from the sample. The
 invention describes a method for producing a streptavidin-binding
 CC protein. This sequence represents a truncation mutant of the streptavidin
 CC -binding peptide SB-19 isolated from Escherichia coli using a method
 CC described in the specification
 XX Sequence 101 AA;

Query Match 82.9%; Score 214; DB 5; Length 101;
 Best Local Similarity 44.6%; Pred. No. 7.7e-22; 0; Mismatches 2; Indels 54; Gaps 1;
 Matches 45; Conservative 2; Mismatches 2;

Qy 1 MDEKTTGWRGGHVGELAGLEQLEQARLLEHHPQGOREPMM----- 40
 Db 1 MDEKTTGWRGGHVGELAGLEQLEQARLLEHHPQGOREP----- 60

Qy 41 ----- SGGCKLG 47
 Db 61 LLDPVVKLTDMWFKKFKNVSKDKRMTPYLEMYDWSGGCKLG 101

RESULT 12

ABU57597
 ID ABU57597 standard; peptide; 101 AA.
 XX
 AC ABU57597;
 XX
 DT 09-APR-2003 (first entry)

DE Streptavidin-binding synthetic peptide SB19 full length peptide.
 XX Streptavidin-binding peptide; mRNA display; peptide library;
 KW fusion protein.
 XX Synthetic.
 OS
 XX US2002155578-A1.
 PN PR
 XX PD 24-OCT-2002.
 XX PP 31-OCT-2001; 2001US-00004381.
 XX
 31-OCT-2000; 2000US-0244541P.
 XX
 PA (SZOSTAK J W.
 PA (WILSON D S.
 PA (KEEFE A D.
 XX
 PI Szostak JW, Wilson DS, Keefe AD;
 XX DR WPI; 2003-182639/18.

XX Novel peptide which binds streptavidin with high affinity useful as
 PT affinity tags for purifying fusion proteins containing proteins of
 PT interest.
 XX
 PS Claim 14; Fig 5; 36pp; English.

XX The invention relates to a peptide which binds streptavidin with a
 CC dissociation constant less than 10 micro M, where the peptide is not
 CC disulphide bonded or cyclised, or with a dissociation constant less than
 CC 10 micro M, where the amino acid sequence of the peptide does not contain
 CC an HPO, HPM, HPN or HOP motif. The peptides are generated by the mRNA
 CC display method to create a library of peptides. Also include are a
 CC nucleic acid encoding the peptide, a vector comprising the nucleic acid,
 CC a fusion protein (and its encoding nucleic acid/vector) comprising a

protein of interest covalently linked to the streptavidin binding peptide and detecting the presence of the fusion protein in a sample using streptavidin or streptavidin-containing compounds. The fusion protein (which comprises the desired protein of interest covalently linked to the streptavidin binding peptide) is useful for purifying a protein of interest from a sample, which involves expressing the fusion protein in the sample, contacting the sample with streptavidin under conditions that allow complex formation between the fusion protein and the streptavidin, isolating the complex and recovering the fusion protein, thereby purifying the protein of interest from the sample. The present sequence is a streptavidin binding peptide of the invention

XX SQ Sequence 101 AA;

Query Match 82.9%; Score 214; DB 6; Length 101;
Best Local Similarity 44.6%; Pred. No. 7.7e-22; Indels 54; Gaps 1;
Matches 45; Conservative 2; Mismatches 0;

QY 1 MDEKTTGWRGGHVGEGLAGELBQLRALKERHHPQGOREPMM----- 40
DB 1 MDEKTTGWRGGHVGEGLAGELBQLRALKERHHPQGOREPMM----- 40

QY 41 -----SGGCKLG 47
DB 61 LLDPEVKUJDWMFKKFKNVSKDCMTFYLEMDWGGCKLG 101

RESULT 13
ID ABG67073 standard; peptide: 101 AA.

XX AC ABG67073;
XX DT 24-SEP-2002 (first entry)

XX DE Streptavidin-binding peptide SB19.
XX KW Streptavidin-binding peptide; mRNA display; peptide library;
XX OS Synthetic.

XX PN US2002155578-A1.
XX PD 24-OCT-2002.

XX PF 31-OCT-2001; 2001US-00004381.

XX PR 31-OCT-2000; 2000US-024441P.

XX PA (SZOS) SZOSTAK, J. W.

PA (WILS) WILSON, D. S.
(KEEF) KEEFE, A. D.

XX PI Szostak JW, Wilson DS,
XX PR WPI; 2003-182639/18.

XX DR Novel peptide which binds streptavidin with high affinity useful as
PT affinity tags for purifying fusion proteins containing proteins of
PT interest.

XX RS Claim 14; Fig 3; 36pp; English.

The invention relates to a peptide which binds streptavidin with a dissociation constant less than 10 micro M, where the peptide is not disulphide bonded or cyclised, or with a dissociation constant less than 10 micro M, where the amino acid sequence of the peptide does not contain an HPO, HPM, HPN or HOP motif. The peptides are generated by the mRNA display method to create a library of peptides. Also include are a nucleic acid encoding the peptide, a vector comprising the nucleic acid, a fusion protein (and its encoding nucleic acid/vector) comprising a protein of interest covalently linked to the streptavidin binding peptide and detecting the presence of the fusion protein in a sample using streptavidin or streptavidin-containing compounds. The fusion protein (which comprises the desired protein of interest covalently linked to the streptavidin binding peptide) is useful for purifying a protein of interest from a sample, which involves expressing the fusion protein in the sample, contacting the sample with streptavidin under conditions that allow complex formation between the fusion protein and the streptavidin, isolating the complex and recovering the fusion protein, thereby purifying the protein of interest from the sample. The present sequence is a Streptavidin binding peptide of the invention

XX SQ Sequence 101 AA;

Query Match 81.8%; Score 211; DB 6; Length 101;
Best Local Similarity 95.0%; Pred. No. 2e-21;

XX SQ Sequence 101 AA;
XX ID ABU57595 standard; peptide: 101 AA.
XX AC ABU57595;
XX DT 09-APR-2003 (first entry)
XX DE Streptavidin-binding synthetic peptide SB19.
XX KW Streptavidin-binding peptide; mRNA display; peptide library;
XX OS Synthetic.

XX PN US2002155578-A1.

XX PD 24-OCT-2002.

XX PF 31-OCT-2001; 2001US-00004381.

XX PR 31-OCT-2000; 2000US-024441P.

XX PA (WILS) WILSON, D. S.
(KEEF) KEEFE, A. D.

XX PI Szostak JW, Wilson DS,
XX PR WPI; 2003-182639/18.

XX DR Novel peptide which binds streptavidin with high affinity useful as
PT affinity tags for purifying fusion proteins containing proteins of
PT interest.

XX RS Claim 14; Fig 3; 36pp; English.

The invention relates to a peptide which binds streptavidin with a dissociation constant less than 10 micro M, where the peptide is not disulphide bonded or cyclised, or with a dissociation constant less than 10 micro M, where the amino acid sequence of the peptide does not contain an HPO, HPM, HPN or HOP motif. The peptides are generated by the mRNA display method to create a library of peptides. Also include are a nucleic acid encoding the peptide, a vector comprising the nucleic acid, a fusion protein (and its encoding nucleic acid/vector) comprising a protein of interest covalently linked to the streptavidin binding peptide and detecting the presence of the fusion protein in a sample using streptavidin or streptavidin-containing compounds. The fusion protein (which comprises the desired protein of interest covalently linked to the streptavidin binding peptide) is useful for purifying a protein of interest from a sample, which involves expressing the fusion protein in the sample, contacting the sample with streptavidin under conditions that allow complex formation between the fusion protein and the streptavidin, isolating the complex and recovering the fusion protein, thereby purifying the protein of interest from the sample. The present sequence is a Streptavidin binding peptide of the invention

XX SQ Sequence 101 AA;

Query Match 81.8%; Score 211; DB 6; Length 101;
Best Local Similarity 95.0%; Pred. No. 2e-21;

Matches 38; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MDEKTTGWRGRGVVEGLAGELBOLARLEHHPOQOREPMM 40
 ID ABG67087 standard; peptide: 38 AA.
 XX
 AC ABG67087;
 XX
 DT 24-SEP-2002 (first entry)
 DE Streptavidin-binding peptide and affinity tag.
 XX
 KW Streptavidin-binding peptide.
 XX
 OS Synthetic.
 XX
 PN WO200238580-A1.
 XX
 PD 16-MAY-2002.
 XX
 PP 31-OCT-2000; 2000WO-US041717.
 XX
 PR 31-OCT-2000; 2000WO-US041717.
 XX
 PA (GEHO) GEN HOSPITAL CORP.
 XX
 PT Novel peptide with high affinity for streptavidin, is expressed as part
 PT of fusion protein to facilitate detection, quantitation and purification
 PT of desired protein.
 XX
 PI Szostak JW, Wilson DS, Keefe AD;
 XX
 DR WPI; 2002-50015/53.
 XX
 PS
 Claim 13; Fig 7A; 56pp; English.
 XX
 CC The invention describes a peptide (I) which binds streptavidin with a
 CC dissociation constant less than 10 nM or 23 nM, where the amino acid
 CC sequence of (I) does not contain an HPO, HPM, HPN or HQP motif, is not
 CC disulfide bonded or cyclised. A fusion protein comprising a protein of
 CC interest covalently linked to (I) is useful for purifying a desired
 CC protein from a sample by contacting the sample with streptavidin under
 CC conditions that allow complex formation between the fusion protein and
 CC the streptavidin, isolating the complex and recovering the fusion
 CC protein, and thus purifying the desired protein from the sample. The
 CC invention describes a method for producing a streptavidin-binding fusion
 CC protein. This sequence represents a streptavidin-binding peptide used in
 CC the creation of streptavidin-binding DNA-tagged peptide described in the
 CC invention.
 XX
 Sequence 38 AA;

Query Match 80.6%; Score 208; DB 5; Length 38;

Best Local Similarity 100.0%; Pred. No. 1.7e-21; Mismatches 0; Indels 0; Gaps 0;
 Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MDEKTTGWRGRGVVEGLAGELBOLARLEHHPOQOREPMM 39
 ID 1 MDEKTTGWRGRGVVEGLAGELBOLARLEHHPOQOREPMM 39
 Db 1 MDEKTTGWRGRGVVEGLAGELBOLARLEHHPOQOREPMM 39

Search completed: July 15, 2004, 20:34:02
 Job time : 55.0824 secs

GenCore version 5.1.6
 Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using SW model

Run on: July 15, 2004, 20:35:51 ; Search time 35:7647 Seconds
 (without alignments)
 332.096 Million cell updates/sec

Title: US-10-004-381-25_COPY_1_38
 perfect score: 208
 Sequence: 1 MDEKTTGWRGGHVVEGLAGELBQLRARLHHHQGQREP 38

Scoring table: BLOSUM62
 Gapop 10.0 , Gapext 0.5

Searched: 1285345 seqs, 31256033 residues

Total number of hits satisfying chosen parameters: 1285345

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
 Listing first 45 summaries

Database : Published Applications At:*

1: /cgmn_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*

2: /cgmn_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*

3: /cgmn_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*

4: /cgmn_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*

5: /cgmn_6/ptodata/1/pubpaa/US07_NEW_PUB_PUB.pep:*

6: /cgmn_6/ptodata/1/pubpaa/PCMVTS_PUBCOMB.pep:*

7: /cgmn_6/ptodata/1/pubpaa/US08_NEW_PUB_PUB.pep:*

8: /cgmn_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*

9: /cgmn_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*

10: /cgmn_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*

11: /cgmn_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*

12: /cgmn_6/ptodata/1/pubpaa/US09_NEW_PUB_PUB.pep:*

13: /cgmn_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*

14: /cgmn_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*

15: /cgmn_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*

16: /cgmn_6/ptodata/1/pubpaa/US10_NEW_PUB_PUB.pep:*

17: /cgmn_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

18: /cgmn_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	208	100.0	38	US-10-004-381-35
2	208	100.0	47	US-10-004-381-25
3	208	100.0	61	US-10-004-381-24
4	208	100.0	75	US-10-004-381-23
5	208	100.0	89	US-10-004-381-22
6	208	100.0	101	US-10-004-381-19
7	208	100.0	101	US-10-004-381-21
8	208	100.0	426	US-10-004-381-41
9	208	100.0	479	US-10-004-381-38
10	193	92.8	47	US-10-004-381-26
11	149	71.6	39	US-10-004-381-27
12	127	61.1	89	US-10-004-381-28
13	102	29	13	US-10-004-381-29
14	67	32.2	102	US-10-004-381-11
15	66.5	32.0	739	US-10-156-761-13097

RESULT 1
 US-10-004-381-35
 ; Sequence 35, Application US/10004381
 ; Publication No. US20020155578A1
 ; GENERAL INFORMATION:
 ; APPLICANT: SZOSTAK, JACK W.
 ; APPLICANT: WILSON, DAVID S.
 ; APPLICANT: KEEFE, ANTHONY D.
 ; TITLE OF INVENTION: STREPTAVIDIN-BINDING PEPTIDES AND USRS
 ; FILE REFERENCE: 00786/388002
 ; CURRENT APPLICATION NUMBER: US/10/004, 381
 ; CURRENT FILING DATE: 2001-10-31
 ; PRIORITY APPLICATION NUMBER: US 60/244,541
 ; PRIORITY FILING DATE: 2000-10-31
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO: 35
 ; LENGTH: 38
 ; TYPE: PRT
 ; ORGANISM: Artificial sequence
 ; FEATURE: Other information: Selected peptide
 ; OTHER INFORMATION: US-10-004-381-35

Query Match 100.0%; Score 208; DB 13; Length 38;
 Best Local Similarity 100.0%; Pred. No. 2. 7e-20;
 Matches 38; Conservative 0; MisMatches 0; Indels 0; Gaps 0;

QY 1 MDEKTTGWRGGHVVEGLAGELBQLRARLHHHQGQREP 38
 DB 1 MDEKTTGWRGGHVVEGLAGELBQLRARLHHHQGQREP 38

RESULT 2
 US-10-004-381-25
 ; Sequence 25, Application US/10004381
 ; Publication No. US20020155578A1

PRIOR FILING DATE: 2000-10-31
 NUMBER OF SEQ ID NOS: 41
 SOFTWARE: FastSEQ for Windows Version 4.0
 SEQ ID NO 19
 LENGTH: 101
 TYPE: PRT
 ORGANISM: Artificial Sequence
 OTHER INFORMATION: selected peptide
 US-10-004-381-19

Query Match 100.0%; Score 208; DB 13; Length 101;
 Best Local Similarity 100.0%; Pred. No. 7.7e-20;
 Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MDEKTTGWRGGHVGVEGLAGELQLQLRLEHHPQGQREP 38
 Db 1 MDEKTTGWRGGHVGVEGLAGELQLQLRLEHHPQGQREP 38

RESULT 7
 US-10-004-381-21
 ; Sequence 21, Application US/10004381
 ; Publication No. US20020155578A1
 ; GENERAL INFORMATION:
 ; APPLICANT: SZOSTAK, JACK W.
 ; APPLICANT: WILSON, DAVID S.
 ; APPLICANT: KEEFE, ANTHONY D.
 ; TITLE OF INVENTION: STREPTAVIDIN-BINDING PEPTIDES AND USES
 ; FILE REFERENCE: 00786/388002
 ; CURRENT APPLICATION NUMBER: US/10/004,381
 ; CURRENT FILING DATE: 2001-10-31
 ; PRIOR APPLICATION NUMBER: US 60/244,541
 ; PRIORITY FILING DATE: 2000-10-31
 ; NUMBER OF SEQ ID NOS: 41
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO 21
 ; LENGTH: 101
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: selected peptide
 US-10-004-381-21

Query Match 100.0%; Score 208; DB 13; Length 101;
 Best Local Similarity 100.0%; Pred. No. 7.7e-20;
 Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MDEKTTGWRGGHVGVEGLAGELQLQLRLEHHPQGQREP 38
 Db 1 MDEKTTGWRGGHVGVEGLAGELQLQLRLEHHPQGQREP 38

RESULT 8
 US-10-004-381-41
 ; Sequence 41, Application US/10004381
 ; Publication No. US20020155578A1
 ; GENERAL INFORMATION:
 ; APPLICANT: SZOSTAK, JACK W.
 ; APPLICANT: WILSON, DAVID S.
 ; APPLICANT: KEEFE, ANTHONY D.
 ; TITLE OF INVENTION: STREPTAVIDIN-BINDING PEPTIDES AND USES
 ; FILE REFERENCE: 00786/388002
 ; CURRENT APPLICATION NUMBER: US/10/004,381
 ; PRIORITY FILING DATE: 2001-10-31
 ; NUMBER OF SEQ ID NOS: 41
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO 41
 ; LENGTH: 426
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: selected peptide
 US-10-004-381-41

Query Match 100.0%; Score 208; DB 13; Length 426;
 Best Local Similarity 100.0%; Pred. No. 3.6e-19;
 Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MDEKTTGWRGGHVGVEGLAGELQLQLRLEHHPQGQREP 38
 Db 1 MDEKTTGWRGGHVGVEGLAGELQLQLRLEHHPQGQREP 415

RESULT 9
 US-10-004-381-38
 ; Sequence 38, Application US/10004381
 ; Publication No. US20020155578A1
 ; GENERAL INFORMATION:
 ; APPLICANT: SZOSTAK, JACK W.
 ; APPLICANT: WILSON, DAVID S.
 ; APPLICANT: KEEFE, ANTHONY D.
 ; TITLE OF INVENTION: STREPTAVIDIN-BINDING PEPTIDES AND USES
 ; FILE REFERENCE: 00786/388002
 ; CURRENT APPLICATION NUMBER: US/10/004,381
 ; CURRENT FILING DATE: 2001-10-31
 ; PRIOR APPLICATION NUMBER: US 60/244,541
 ; PRIORITY FILING DATE: 2000-10-31
 ; NUMBER OF SEQ ID NOS: 41
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO 38
 ; LENGTH: 479
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Designed protein
 US-10-004-381-38

Query Match 100.0%; Score 208; DB 13; Length 479;
 Best Local Similarity 100.0%; Pred. No. 4.1e-19;
 Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MDEKTTGWRGGHVGVEGLAGELQLQLRLEHHPQGQREP 38
 Db 1 MDEKTTGWRGGHVGVEGLAGELQLQLRLEHHPQGQREP 415

RESULT 10
 US-10-004-381-26
 ; Sequence 26, Application US/10004381
 ; Publication No. US20020155578A1
 ; GENERAL INFORMATION:
 ; APPLICANT: SZOSTAK, JACK W.
 ; APPLICANT: WILSON, DAVID S.
 ; APPLICANT: KEEFE, ANTHONY D.
 ; TITLE OF INVENTION: STREPTAVIDIN-BINDING PEPTIDES AND USES
 ; FILE REFERENCE: 00786/388002
 ; CURRENT APPLICATION NUMBER: US/10/004,381
 ; CURRENT FILING DATE: 2001-10-31
 ; PRIOR APPLICATION NUMBER: US 60/244,541
 ; PRIORITY FILING DATE: 2000-10-31
 ; NUMBER OF SEQ ID NOS: 41
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO 26
 ; LENGTH: 47
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: selected peptide
 US-10-004-381-26

Query Match 92.8%; Score 193; DB 13; Length 47;
 Best Local Similarity 94.7%; Pred. No. 3.3e-18; Mismatches 2;
 Matches 36; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MDEKTTGWRGHHVEGLAGELBOLARLEHHPQOREP 38
 Db 1 MDEKTTGWRGHHVEGLAGELBOLARLEHHPQOREP 38

RESULT 11
 US-10-004-381-27
 Sequence 27, Application US/10004381
 Publication No. US20020155578A1
 GENERAL INFORMATION:
 APPLICANT: SZOSTAK, JACK W.
 APPLICANT: WILSON, DAVID S.
 APPLICANT: KEEFE, ANTHONY D.
 TITLE OF INVENTION: STREPTAVIDIN-BINDING PEPTIDES AND USES
 TITLE OF INVENTION: THEREOF
 FILE REFERENCE: 00786/388002
 CURRENT APPLICATION NUMBER: US/10/004, 381
 CURRENT FILING DATE: 2001-10-31
 PRIOR APPLICATION NUMBER: US 60/244,541
 PRIOR FILING DATE: 2000-10-31
 NUMBER OF SEQ ID NOS: 41
 SOFTWARE: FastSEQ for Windows Version 4.0
 SEQ ID NO 27
 LENGTH: 39
 TYPE: PRT
 ORGANISM: Artificial Sequence
 FEATURE: OTHER INFORMATION: selected peptide
 US-10-004-381-27

Query Match 71.6%; Score 149; DB 13; Length 39;
 Best Local Similarity 100.0%; Pred. No. 1.9e-12;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 GRVVEGLAGELBOLARLEHHPQOREP 38
 Db 3 GRVVEGLAGELBOLARLEHHPQOREP 30

RESULT 12
 US-10-004-381-28
 Sequence 28, Application US/10004381
 Publication No. US20020155578A1
 GENERAL INFORMATION:
 APPLICANT: SZOSTAK, JACK W.
 APPLICANT: WILSON, DAVID S.
 APPLICANT: KEEFE, ANTHONY D.
 TITLE OF INVENTION: STREPTAVIDIN-BINDING PEPTIDES AND USES
 TITLE OF INVENTION: THEREOF
 FILE REFERENCE: 00786/388002
 CURRENT APPLICATION NUMBER: US/10/004, 381
 CURRENT FILING DATE: 2001-10-31
 PRIOR APPLICATION NUMBER: US 60/244,541
 PRIOR FILING DATE: 2000-10-31
 NUMBER OF SEQ ID NOS: 41
 SOFTWARE: FastSEQ for Windows Version 4.0
 SEQ ID NO 28
 LENGTH: 89
 TYPE: PRT
 ORGANISM: Artificial Sequence
 FEATURE: OTHER INFORMATION: selected peptide
 US-10-004-381-28

Query Match 61.1%; Score 127; DB 13; Length 89;
 Best Local Similarity 100.0%; Pred. No. 3.9e-09;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 EGLAGELBOLARLEHHPQOREP 38
 Db 3 EGLAGELBOLARLEHHPQOREP 26

RESULT 13
 US-10-004-381-29
 Sequence 29, Application US/10004381
 Publication No. US20020155578A1
 GENERAL INFORMATION:
 APPLICANT: SZOSTAK, JACK W.
 APPLICANT: WILSON, DAVID S.
 APPLICANT: KEEFE, ANTHONY D.
 TITLE OF INVENTION: STREPTAVIDIN-BINDING PEPTIDES AND USES
 TITLE OF INVENTION: THEREOF
 FILE REFERENCE: 00786/388002
 CURRENT APPLICATION NUMBER: US/10/004, 381
 CURRENT FILING DATE: 2001-10-31
 PRIOR APPLICATION NUMBER: US 60/244,541
 PRIOR FILING DATE: 2000-10-31
 NUMBER OF SEQ ID NOS: 41
 SOFTWARE: FastSEQ for Windows Version 4.0
 SEQ ID NO 29
 LENGTH: 29
 TYPE: PRT
 ORGANISM: Artificial Sequence
 FEATURE: OTHER INFORMATION: selected peptide
 US-10-004-381-29

Query Match 49.0%; Score 102; DB 13; Length 29;
 Best Local Similarity 100.0%; Pred. No. 2.4e-06;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 ELEQRLARLEHHPQOREP 38
 Db 2 ELEQRLARLASHHPQOREP 20

RESULT 14
 US-10-004-381-11
 Sequence 11, Application US/10004381
 Publication No. US20020155578A1
 GENERAL INFORMATION:
 APPLICANT: SZOSTAK, JACK W.
 APPLICANT: WILSON, DAVID S.
 APPLICANT: KEEFE, ANTHONY D.
 TITLE OF INVENTION: STREPTAVIDIN-BINDING PEPTIDES AND USES
 TITLE OF INVENTION: THEREOF
 FILE REFERENCE: 00786/388002
 CURRENT APPLICATION NUMBER: US/10/004, 381
 CURRENT FILING DATE: 2001-10-31
 PRIOR APPLICATION NUMBER: US 60/244,541
 PRIOR FILING DATE: 2000-10-31
 NUMBER OF SEQ ID NOS: 41
 SOFTWARE: FastSEQ for Windows Version 4.0
 SEQ ID NO 11
 LENGTH: 102
 TYPE: PRT
 ORGANISM: Artificial Sequence
 FEATURE: OTHER INFORMATION: selected peptide
 US-10-004-381-11

Query Match 32.2%; Score 67; DB 13; Length 102;
 Best Local Similarity 45.0%; Pred. No. 0.42;
 Matches 18; Conservative 2; Mismatches 6; Indels 14; Gaps 3;

Qy 1 MDEKTTGWRGHHVEGLAGELBOLARLEHHPQOREP 38
 Db 1 MDEK-TWHERVHHLAD-----LEQHQGQR 28

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2004, 20:29:44 ; Search time 42.9176 Seconds

(without alignments)
250.172 Million cell updates/sec

Title: US-10-004-381-25_COPY_1_38

Perfect score: 208
Sequence: 1 MDEKTTGWRQGHVVEGLAGBLEQLRARLLEHHPOQOREP 38

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

A_Geneseq_29Jan04:*

1: geneseqD1980s:*

2: geneseqD1990s:*

3: geneseqD2000s:*

4: geneseqD2001s:*

5: geneseqD2002s:*

6: geneseqD2003as:*

7: geneseqD2003bs:*

geneseqD2004as:*

SUMMARIES

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

ALIGNMENTS

RESULT	1
ID	ABG67087
XX	standard; peptide; 38 AA.
AC	ABG67087;
XX	
DT	24-SEP-2002 (first entry)
XX	
DE	Streptavidin-binding peptide and affinity tag.
XX	
KW	Streptavidin-binding peptide.
XX	
OS	Synthetic.
XX	
PN	WO200338580-A1.
XX	
PD	16-MAY-2002.
XX	
PF	31-OCT-2000; 2000WO-US041717.
XX	
PR	31-OCT-2000; 2000WO-US041717.
XX	
PA	(GEHO) GEN HOSPITAL CORP.
XX	
PI	Szostak JW, Wilson DS, Keefe AD;
XX	
DR	WPI; 2002-500155/53.
XX	
PT	Novel peptide with high affinity for streptavidin, is expressed as part of fusion protein to facilitate detection, quantitation and purification of desired protein.
XX	
PS	Claim 13; Fig 7A; 56pp; English.
XX	
CC	The invention describes a peptide (I) which binds streptavidin with a dissociation constant less than 10 nM or 23 nM, where the amino acid sequence of (I) does not contain an HPQ, HPM, HPN or HQP motif, is not disulfide bonded or cyclized. A fusion protein comprising a protein of interest covalently linked to (I) is useful for purifying a desired protein from a sample by contacting the sample with streptavidin under conditions that allow complex formation between the fusion protein and the streptavidin, isolating the complex and recovering the fusion protein, and thus purifying the desired protein from the sample. The invention describes a method for producing a streptavidin-binding fusion protein. This sequence represents a streptavidin-binding peptide used in the creation of streptavidin-binding DNA-tagged peptide described in the invention.
CC	

Abu57605 Synthetic
Abg67065 Streptavi
Ab57587 Streptavi
Abg67066 Streptavi
Abu57588 Streptavi
Abu63619 Human gas
Abg25060 Novel hum
Abp09375 Human ORF
Abg6152 Human Nov
Abg6153 FRAP homo
Abg6149 Human Nov
Abg96159 Human Nov
Abg96160 Human Nov
Abg96162 Human Nov

Result No.	Score	Query Match Length	DB ID	Description
1	208	100.0	38	5 ABG67087
2	208	100.0	38	6 ABU7609
3	208	100.0	47	5 ABG67079
4	208	100.0	47	6 ABU57601
5	208	100.0	61	5 ABG67078
6	208	100.0	61	6 ABU57600
7	208	100.0	75	5 ABG67077
8	208	100.0	75	6 ABU57599
9	208	100.0	89	5 ABG67076
10	208	100.0	89	6 ABU57598
11	208	100.0	101	5 ABG67075
12	208	100.0	101	5 ABG67073
13	208	100.0	101	6 ABU57597
14	208	100.0	101	6 ABU57595
15	208	100.0	426	5 ABG67091
16	208	100.0	426	6 ABU57613
17	208	100.0	479	5 ABG67089
18	208	100.0	479	5 ABG67611
19	208	92.8	47	5 ABG67080
20	193	92.8	47	6 ABU57602
21	149	71.6	39	5 ABG67081
22	149	71.5	39	6 ABU57603
23	127	61.1	89	5 ABG67082
24	127	61.1	89	5 ABG67064 Synthetic
25	102	49.0	49.0	5 ABG67083 Streptavi

us-10-004-381-25_copy_1_38.rag

SQ	Sequence 38 AA;	Matches	38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Query	MDBKTTGRRGGHVGEGLAGELQLRARLEHHPOGOREP 38	1	MDBKTTGRRGGHVGEGLAGELQLRARLEHHPOGOREP 38
Best Local Similarity	100.0%; Pred. No. 4.1e-22; Mismatches 0; Indels 0; Gaps 0;		
Matches	38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
ABU57609	MDBKTTGRRGGHVGEGLAGELQLRARLEHHPOGOREP 38	1	MDBKTTGRRGGHVGEGLAGELQLRARLEHHPOGOREP 38
ID	ABU57609 standard; peptide; 38 AA.		
XX			
AC	ABU57609;		
XX			
DT	09-APR-2003 (first entry)		
XX			
DE	Streptavidin-binding synthetic peptide SB19 C4 N-terminal affinity tag.		
XX			
KW	Streptavidin-binding peptide; mRNA display; peptide library; fusion protein; SB19.		
XX			
OS	Synthetic.		
XX			
DN	US2002155578-A1.		
XX			
PD	24-OCT-2002.		
XX			
PF	31-OCT-2001; 2001US-00004381.		
XX			
PR	31-OCT-2000; 2000US-0244541P.		
XX			
PA	(SZOS/) SZOSTAK JW.		
XX			
PA	(WILS/) WILSON DS.		
XX			
PA	(KEEF/) KEEFE AD.		
XX			
PR	Szostak JW, Wilson DS, Keefe AD;		
XX			
PT	WPI; 2003-182639/18.		
XX			
PR	Novel peptide which binds streptavidin with high affinity useful as affinity tag for purifying fusion proteins containing proteins of interest.		
XX			
PS	Claim 14; FIG 7A; 36PP; English.		
XX			
CC	The invention relates to a peptide which binds streptavidin with a dissociation constant less than 10 micro M, where the peptide is not disulfide bonded or cyclized, or with a dissociation constant less than 10 micro M, where the amino acid sequence of the peptide does not contain an HPQ, HPM, HPN or HQP motif. The peptides are generated by the mRNA display method to create a library of peptides. Also include are a nucleic acid encoding the peptide, a vector comprising the nucleic acid, a fusion protein (and its encoding nucleic acid/vector) comprising a protein of interest covalently linked to the streptavidin binding peptide and detecting the presence of the fusion protein in a sample using streptavidin or streptavidin-containing compounds. The fusion protein (which comprises the streptavidin-containing protein of interest covalently linked to the streptavidin binding peptide) is useful for purifying a protein of interest from a sample, which involves expressing the fusion protein in the sample, contacting the sample with streptavidin under conditions that allow complex formation between the fusion protein and the streptavidin, isolating the complex and recovering the fusion protein, and thus purifying the desired protein from the sample. The present invention describes a method for producing a streptavidin-binding peptide SB-19 isolated from Escherichia coli using a method described in the specification		
CC	Sequence 47 AA;		
XX			
Query	MDBKTTGRRGGHVGEGLAGELQLRARLEHHPOGOREP 38	1	MDBKTTGRRGGHVGEGLAGELQLRARLEHHPOGOREP 38
Best Local Similarity	100.0%; Pred. No. 5.2e-22; Mismatches 0; Indels 0; Gaps 0;		
Matches	38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
ABG67079	MDBKTTGRRGGHVGEGLAGELQLRARLEHHPOGOREP 38	1	MDBKTTGRRGGHVGEGLAGELQLRARLEHHPOGOREP 38
ID	ABG67079 standard; peptide; 47 AA.		
XX			
AC	ABG67079;		
XX			
DT	24-SEP-2002 (first entry)		
XX			
DE	Streptavidin-binding peptide SB19 truncation mutant #5.		
XX			
OS	Escherichia coli.		
XX			
PN	WO200238580-A1.		
XX			
PD	16-MAY-2002.		
XX			
PF	31-OCT-2001; 2000WO-US041717.		
XX			
PR	31-OCT-2000; 2000WO-US041717.		
XX			
PA	(GEHO) GEN HOSPITAL CORP.		
XX			
PI	Szostak JW, Wilson DS, Keefe AD;		
XX			
DR	WPI; 2002-500155/53.		
XX			
XX	Novel peptide with high affinity for streptavidin, is expressed as part of fusion protein to facilitate detection, quantitation and purification of desired protein.		
XX			
PS	Claim 9; FIG 5; 56PP; English.		
XX			
CC	The invention describes a peptide (I) which binds streptavidin with a dissociation constant less than 10 mM or 23 nM, where the amino acid sequence of (I) does not contain an HPQ, HPM, HPN or HQP motif, is not disulfide bonded or cyclized. A fusion protein comprising protein of interest covalently linked to (I) is useful for purifying a desired protein from a sample by contacting the sample with streptavidin under conditions that allow complex formation between the fusion protein and the streptavidin, isolating the complex and recovering the fusion protein, and thus purifying the desired protein from the sample. The invention describes a method for producing a streptavidin-binding fusion protein. This sequence represents a truncation mutant of the streptavidin-binding peptide SB-19 isolated from Escherichia coli using a method described in the specification		
CC	Sequence 47 AA;		
XX			
Query	MDBKTTGRRGGHVGEGLAGELQLRARLEHHPOGOREP 38	1	MDBKTTGRRGGHVGEGLAGELQLRARLEHHPOGOREP 38
Best Local Similarity	100.0%; Pred. No. 5.2e-22; Mismatches 0; Indels 0; Gaps 0;		
Matches	38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
ABU57601	MDBKTTGRRGGHVGEGLAGELQLRARLEHHPOGOREP 38	1	MDBKTTGRRGGHVGEGLAGELQLRARLEHHPOGOREP 38
ID	ABU57601 standard; peptide; 47 AA.		
XX			
Sequence	38 AA;		
SQ			

XX	Streptavidin-binding peptide; mutant; mutein.
DB	Synthetic peptide SB19 C-terminal deletion mutant, C4.
KW	Streptavidin-binding peptide; mRNA display; peptide library;
KW	fusion protein; mutant; mutein; SB19.
KW	Synthetic.
OS	
XX	
PN	US2002155578-A1.
XX	
PD	24-OCT-2002.
XX	
PR	31-OCT-2000; 2000WO-US041717.
XX	
PR	31-OCT-2000; 2000WO-US041717.
XX	
PA	(GEHO) GEN HOSPITAL CORP.
XX	
PR	Szostak JW, Wilson DS, Keefe AD;
XX	
PA	(SZOS/) SZOSTAK J W.
XX	
PT	31-OCT-2000; 2000US-0244541P.
XX	
PA	(WILS/) WILSON D S.
PA	(KEEF/) KEEFE A D.
XX	
DR	WPI; 2003-182639/18.
XX	
PT	Novel peptide which binds streptavidin with high affinity useful as
PT	affinity tags for purifying fusion proteins containing proteins of
PT	interest.
XX	
PS	Claim 14; Fig 5; 36pp; English.
XX	
CC	The invention relates to a peptide which binds streptavidin with a
CC	dissociation constant less than 10 micro M, where the peptide is not
CC	disulfide bonded or cyclized, or with a dissociation constant less than
CC	10 micro M, where the amino acid sequence of the peptide does not contain
CC	an HPO, HPM, HPN or HOP motif. The peptides are generated by the mRNA
CC	display method to create a library of peptides. Also include are a
CC	nucleic acid encoding the peptide, a vector comprising the nucleic acid,
CC	a fusion protein (and its encoding nucleic acid/vector) comprising a
CC	protein of interest covalently linked to the streptavidin binding peptide
CC	and detecting the presence of the fusion protein in a sample using
CC	streptavidin or streptavidin-containing compounds. The fusion protein
CC	(which comprises the desired protein of interest covalently linked to the
CC	streptavidin binding peptide) is useful for purifying a protein of
CC	interest from a sample, which involves expressing the fusion protein in
CC	the sample, contacting the sample with streptavidin under conditions that
CC	allow complex formation between the fusion protein and the streptavidin,
CC	isolating the complex and recovering the fusion protein, thereby
CC	purifying the protein of interest from the sample. The present sequence
CC	is a mutated version of SB19, a streptavidin binding peptide of the
CC	invention. The mutants comprise 3 N-terminal deletions, 3 C-terminal
CC	deletions and a truncated N-terminal peptide with point mutations
XX	
SQ	Sequence 47 AA;
Query Match	100 %; Score 208; DB 5; Length 61;
Best Local Similarity	100 %; Pred. No. 7e-22;
Matches	38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy	1 MDEKTTGWRGGHVEGLAGELBOLRARLEHHPQGOREP 38
Db	1 MDEKTTGWRGGHVEGLAGELBOLRARLEHHPQGOREP 38
RESULT 6	
ID	ABU57600
XX	
AC	ABU57600;
XX	
AC	ABU57600;
XX	
DT	09-APR-2003 (first entry)
XX	
DE	Synthetic peptide SB19 C-terminal deletion mutant, C3.
XX	
KW	Streptavidin-binding peptide; mRNA display; peptide library;
KW	fusion protein; mutant; mutein; SB19.
OS	Synthetic.
XX	
PN	US2002155578-A1.
XX	
PD	24-OCT-2002.
XX	
PR	31-OCT-2001; 2001US-00004381.
XX	
PR	31-OCT-2000; 2000US-0244541P.
XX	
PA	(SZOS/) SZOSTAK J W.
XX	
RESULT 5	
ID	ABG67078
AC	ABG67078;
XX	
DT	24-SEP-2002 (first entry)
XX	
DB	Streptavidin-binding peptide SB19 truncation mutant #4.
XX	

PA (WILS/) WILSON D S.
 PA (KEEF/) KEEFE A D.
 XX
 PT
 PT
 PT
 PT
 XX
 DR
 WPI; 2003-182639/18.

XX
 PT Novel peptide which binds streptavidin with high affinity useful as
 PT affinity tags for purifying fusion proteins containing proteins of
 PT interest.

XX
 PS
 Claim 14; Fig 5; 36pp; English.

XX
 The invention relates to a peptide which binds streptavidin with a
 CC dissociation constant less than 10 micro M, where the peptide is not
 CC disulphide bonded or cyclised, or with a dissociation constant less than
 CC 10 micro M, where the amino acid sequence of the peptide does not contain
 CC an HPO, HPM, HPN or HOP motif. The peptides are generated by the mRNA
 CC display method to create a library of peptides. Also include are a
 CC nucleic acid encoding the peptide, a vector comprising the nucleic acid,
 CC a fusion protein (and its encoding nucleic acid/vector) comprising a
 CC protein of interest covalently linked to the streptavidin binding peptide
 CC and detecting the presence of the fusion protein in a sample using
 CC streptavidin or streptavidin-containing compounds. The fusion protein
 (which comprises the desired protein of interest covalently linked to the
 CC streptavidin binding peptide) is useful for purifying a protein of
 CC interest from a sample, which involves expressing the fusion protein in
 CC the sample, contacting the sample with streptavidin under conditions that
 CC allow complex formation between the fusion protein and the streptavidin,
 CC isolating the complex and recovering the fusion protein, thereby
 CC purifying the protein of interest from the sample. The present sequence
 CC is a mutated version of SB19, a streptavidin binding peptide of the
 CC invention. The mutants comprise 3 N-terminal deletions, 3 C-terminal
 CC deletions and a truncated N-terminal peptide with point mutations
 XX
 SQ Sequence 61 AA;

Query Match 100.0%; Score 208; DB 6; Length 61;
 Best Local Similarity 100.0%; Pred. No. 7e-22;
 Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QV 1 MDEKTTGWRGGHVGEGLAGELQARLLEHHRPOQREP 38
 Db 1 MDEKTTGWRGGHVGEGLAGELQARLLEHHRPOQREP 38

RESULT 7
 ABG67077
 ID ABG67077 standard; peptide; 75 AA.
 XX
 AC ABG67077;
 XX
 DT 24-SEP-2002 (first entry)
 XX
 DB Streptavidin-binding peptide SB19 truncation mutant #3.
 KW Streptavidin-binding peptide; mRNA display; peptide library;
 KW fusion protein; mutant; mutein; SB19.
 OS Synthetic.
 XX
 DN US2002155578-A1.
 XX
 PD 24-OCT-2002.
 XX
 PF 31-OCT-2001; 2001US-00004381.
 XX
 PR 31-OCT-2000; 2000US-0244541P.
 XX
 PA (SZOS/) SZOSTAK J W.
 PA (WILS/) WILSON D S.
 PA (KEEF/) KEEFE A D.
 XX
 PT Szostak JW, Wilson DS, KEEFE AD;
 XX
 PN WO200238580-A1.
 XX
 PD 16-MAY-2002.
 XX
 PR 31-OCT-2000; 2000WO-US041717.
 XX
 PR 31-OCT-2000; 2000WO-US041717.
 XX
 PA (GEHO) GEN HOSPITAL CORP.
 XX
 PT Szostak JW, Wilson DS, KEEFE AD;
 XX
 DR WPI; 2002-500155/53.

XX
 PS
 Claim 14; Fig 5; 36pp; English.

XX
 PT Novel peptide which binds streptavidin with high affinity useful as
 PT affinity tags for purifying fusion proteins containing proteins of
 PT interest.

XX
 The invention relates to a peptide which binds streptavidin with a
 CC dissociation constant less than 10 micro M, where the peptide is not
 CC disulphide bonded or cyclised, or with a dissociation constant less than
 CC 10 micro M, where the amino acid sequence of the peptide does not contain

CC an HPO, HPM, HPN or HQP motif. The peptides are generated by the mRNA
 CC display method to create a library of peptides. Also include are a
 CC nucleic acid encoding the peptide, a vector comprising the nucleic acid,
 CC a fusion protein (and its encoding nucleic acid/vector) comprising a
 CC protein of interest covalently linked to the streptavidin binding peptide
 CC and detecting the presence of the fusion protein in a sample using
 CC streptavidin or streptavidin-containing compounds. The fusion protein
 CC (which comprises the desired protein of interest covalently linked to the
 CC streptavidin binding peptide) is useful for purifying a protein of
 CC interest from a sample, which involves expressing the fusion protein in
 CC the sample, contacting the sample with streptavidin under conditions that
 CC allow complex formation between the fusion protein and the streptavidin,
 CC isolating the complex and recovering the fusion protein, thereby
 CC purifying the protein of interest from the sample. The present sequence
 CC is a mutated version of SB19, a streptavidin binding peptide of the
 CC invention. The mutants comprise 3 N-terminal deletions, 3 C-terminal
 CC deletions and a truncated N-terminal peptide with point mutations
 XX

SQ Sequence 75 AA;

Query Match 100.0%; Score 208; DB 6;
 Best Local Similarity 100.0%; Pred. No. 8. 8e-22;
 Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MDEKTTGMRGHHVVEGLAGELBEQARLARLHHPQOREP 38
 Db 1 MDEKTTGMRGHHVVEGLAGELBEQARLARLHHPQOREP 38

RESULT 9

ABG67076 ID ABG67076 standard; peptide; 89 AA.

XX

AC ABG67076;

XX

DT 24-SEP-2002 (first entry)

XX

DE Streptavidin-binding Peptide SB19 truncation mutant #2.

XX

KW Streptavidin-binding Peptide; mutant; mutein.

XX

OS Escherichia coli.

XX

PN WO200238580-A1.

XX

PD 16-MAY-2002.

XX

PR 31-OCT-2000; 2000WO-US041717.

XX

PR 31-OCT-2000; 2000WO-US041717.

XX

PA (GEHO) GEN HOSPITAL CORP.

XX

PT Szostak JW, Wilson DS, Keefe AD;

XX

PS DR; WPI; 2003-182639/18.

XX

PT Novel peptide which binds streptavidin with high affinity useful as
 PR affinity tag for purifying fusion proteins containing proteins of
 PR interest.

XX

PS Claim 14; Fig 5; 36pp; English.

XX

CC The invention relates to a peptide which binds streptavidin with a
 CC dissociation constant less than 10 micro M, where the peptide is not
 CC disulphide bonded or cyclised, or with a dissociation constant less than
 CC 10 micro M, where the amino acid sequence of the peptide does not contain
 CC an HPO, HPM, HPN or HQP motif. The peptides are generated by the mRNA
 CC display method to create a library of peptides. Also include are a
 CC nucleic acid encoding the peptide, a vector comprising the nucleic acid,
 CC a fusion protein (and its encoding nucleic acid/vector) comprising a
 CC protein of interest covalently linked to the streptavidin binding peptide
 CC and detecting the presence of the fusion protein in a sample using
 CC streptavidin or streptavidin-containing compounds. The fusion protein
 CC (which comprises the desired protein of interest covalently linked to the
 CC streptavidin binding peptide) is useful for purifying a protein of
 CC interest from a sample, which involves expressing the fusion protein in
 CC the sample, contacting the sample with streptavidin under conditions that
 CC allow complex formation between the fusion protein and the streptavidin,
 CC isolating the complex and recovering the fusion protein, thereby
 CC purifying the protein of interest from the sample. The present sequence
 CC is a mutated version of SB19, a streptavidin binding peptide of the
 CC invention. The mutants comprise 3 N-terminal deletions, 3 C-terminal
 CC deletions and a truncated N-terminal peptide with point mutations

CC protein. This sequence represents a truncation mutant of the streptavidin
 CC -binding peptide SB19 isolated from Escherichia coli using a method
 CC described in the specification
 XX

SQ Sequence 89 AA;

Query Match 100.0%; Score 208; DB 5; Length 89;
 Best Local Similarity 100.0%; Pred. No. 1.1e-21;
 Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDEKTTGMRGHHVVEGLAGELBEQARLARLHHPQOREP 38
 Db 1 MDEKTTGMRGHHVVEGLAGELBEQARLARLHHPQOREP 38

RESULT 10

ABU5798 ID ABU5798 standard; peptide; 89 AA.

XX

AC ABU5798;

XX

DT 09-APR-2003 (first entry)

XX

DE Synthetic peptide SB19 C-terminal deletion mutant, Cl.

XX

KW Streptavidin-binding Peptide; mRNA display; peptide library;

XX

OS Synthetic.

XX

PN US2002155578-A1.

XX

PD 24-OCT-2002.

XX

PR 31-OCT-2001; 2001US-00000381.

XX

PR 31-OCT-2000; 2000US-024441P.

XX

PA (SZOS/) SZOSTAK J W.

PA (WILS/) WILSON D S.

PA (KEEF/) KEEFE A D.

XX

PI Szostak JW, Wilson DS, Keefe AD;

XX

PS DR; WPI; 2003-182639/18.

XX

PT Novel peptide which binds streptavidin with high affinity useful as
 PR affinity tag for purifying fusion proteins containing proteins of
 PR interest.

XX

PS Claim 14; Fig 5; 36pp; English.

XX

CC The invention relates to a peptide which binds streptavidin with a
 CC dissociation constant less than 10 micro M, where the peptide is not
 CC disulphide bonded or cyclised, or with a dissociation constant less than
 CC 10 micro M, where the amino acid sequence of the peptide does not contain
 CC an HPO, HPM, HPN or HQP motif. The peptides are generated by the mRNA
 CC display method to create a library of peptides. Also include are a
 CC nucleic acid encoding the peptide, a vector comprising the nucleic acid,
 CC a fusion protein (and its encoding nucleic acid/vector) comprising a
 CC protein of interest covalently linked to the streptavidin binding peptide
 CC and detecting the presence of the fusion protein in a sample using
 CC streptavidin or streptavidin-containing compounds. The fusion protein
 CC (which comprises the desired protein of interest covalently linked to the
 CC streptavidin binding peptide) is useful for purifying a protein of
 CC interest from a sample, which involves expressing the fusion protein in
 CC the sample, contacting the sample with streptavidin under conditions that
 CC allow complex formation between the fusion protein and the streptavidin,
 CC isolating the complex and recovering the fusion protein, thereby
 CC purifying the protein of interest from the sample. The present sequence
 CC is a mutated version of SB19, a streptavidin binding peptide of the
 CC invention. The mutants comprise 3 N-terminal deletions, 3 C-terminal
 CC deletions and a truncated N-terminal peptide with point mutations

XK
SQ Sequence 89 AA;

Query Match 100.0%; Score 208; DB 6; Length 89;

Best Local Similarity 100.0%; Pred. No. 1.1e-21;

Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MDEKTTGWRGAGHVGELAEQQLARLEHHPOQREP 38

Db 1 MDEKTTGWRGAGHVGELAEQQLARLEHHPOQREP 38

RESULT 11
ABG67075 standard; peptide: 101 AA.

XX
ID ABG67075 standard; peptide: 101 AA.

XX
ID ABG67073;

XX
AC ABG67073;

XX
DT 24-SEP-2002 (first entry)

XX
DE Streptavidin-binding peptide SB19.

XX
KW Streptavidin-binding peptide.

XX
OS Escherichia coli.

XX
BN WO200238580-A1.

XX
ID 16-MAY-2002.

XX
PR 31-OCT-2000; 2000WO-US041717.

XX
PA (GEHO) GEN HOSPITAL CORP.

XX
PI Szostak JW, Wilson DS, Keefe AD;

XX
DR WPI; 2002-500155/53.

XX
XN Novel peptide with high affinity for streptavidin, is expressed as part

XX
PT of fusion protein to facilitate detection, quantitation and purification

XX
PT of desired protein.

XX
PS Claim 9; Fig 3; 56pp; English.

XX
CC The invention describes a peptide (I) which binds streptavidin with a
CC dissociation constant less than 10 μ M or 23 nM, where the amino acid
CC sequence of (I) does not contain an HPO, HPM, HPN or HOP motif, is not
CC disulfide bonded or cyclised. A fusion protein comprising a protein of
CC interest covalently linked to (I) is useful for purifying a desired
CC protein from a sample by contacting the sample with streptavidin under
CC conditions that allow complex formation between the fusion protein and
CC the streptavidin, isolating the complex and recovering the fusion
CC protein, and thus purifying the desired protein from the sample. The
CC invention describes a method for producing a streptavidin-binding fusion
CC protein. This sequence represents a streptavidin-binding peptide isolated
CC from Escherichia coli using a method described in the specification

XX
SQ The invention describes a peptide (I) which binds streptavidin with a
CC dissociation constant less than 10 μ M or 23 nM, where the amino acid
CC sequence of (I) does not contain an HPO, HPM, HPN or HOP motif, is not
CC disulfide bonded or cyclised. A fusion protein comprising a protein of
CC interest covalently linked to (I) is useful for purifying a desired
CC protein from a sample by contacting the sample with streptavidin under
CC conditions that allow complex formation between the fusion protein and
CC the streptavidin, isolating the complex and recovering the fusion
CC protein, and thus purifying the desired protein from the sample. The
CC invention describes a method for producing a streptavidin-binding fusion
CC protein. This sequence represents a streptavidin-binding peptide isolated
CC from Escherichia coli using a method described in the specification

XX
SQ Sequence 101 AA;

Query Match 100.0%; Score 208; DB 5; Length 101;

Best Local Similarity 100.0%; Pred. No. 1.2e-21;

Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MDEKTTGWRGAGHVGELAEQQLARLEHHPOQREP 38

Db 1 MDEKTTGWRGAGHVGELAEQQLARLEHHPOQREP 38

RESULT 13
ABU57597 standard; peptide: 101 AA.

XX
ID ABU57597

XX
AC ABU57597;

XX
DT 09-APR-2003 (first entry)

XX
DE Streptavidin-binding synthetic peptide SB19 full length peptide.

XX
KW Streptavidin-binding peptide; mRNA display; peptide library;

XX
OS Synthetic.

XX
PN US2002155578-A1.

RESULT 12

RESULT 12

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

Om protein - protein search, using sw model

Run on: July 15, 2004, 20:31:20 ; Search time 12.7176 Seconds
(without alignments)
355.490 Million cell updates/sec

Title: US-10-004-381-25
perfect score: 258
Sequence: 1 MDEKITGWRQGHVVEGLAGE.....LEHHHQGOREPMMGGCKLG 47

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78;*
1: pir1;*
2: pir2;*
3: pir3;*
4: pir4;*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No. Score Match Length DB ID Description

1 73.5 28.5 554 2 B87375 conserved hypothetical protein CC1017 [imported] - Caulobacter crescentus

2 65.5 25.4 528 2 C87663 conserved hypothetical protein CC1017 [imported] - Caulobacter crescentus

3 61.2 23.6 429 2 F48015 conserved hypothetical protein CC1017 [imported] - Caulobacter crescentus

4 59.5 23.1 342 2 E64302 conserved hypothetical protein CC1017 [imported] - Caulobacter crescentus

5 59.5 23.1 342 2 A90894 conserved hypothetical protein CC1017 [imported] - Caulobacter crescentus

6 59.5 23.1 342 2 H85723 conserved hypothetical protein CC1017 [imported] - Caulobacter crescentus

7 59.5 22.9 1974 2 T300101 conserved hypothetical protein CC1017 [imported] - Caulobacter crescentus

8 58.5 22.7 302 2 B82189 conserved hypothetical protein CC1017 [imported] - Caulobacter crescentus

9 57.0 22.1 474 1 E24723 conserved hypothetical protein CC1017 [imported] - Caulobacter crescentus

10 57.0 22.1 496 2 H70068 conserved hypothetical protein CC1017 [imported] - Caulobacter crescentus

11 57.0 22.1 663 2 D83545 conserved hypothetical protein CC1017 [imported] - Caulobacter crescentus

12 56.5 21.9 215 2 A11074 conserved hypothetical protein CC1017 [imported] - Caulobacter crescentus

13 56.5 21.9 1560 2 I54361 conserved hypothetical protein CC1017 [imported] - Caulobacter crescentus

14 56.0 21.7 261 2 S43218 conserved hypothetical protein CC1017 [imported] - Caulobacter crescentus

15 56.0 21.7 338 2 T36025 conserved hypothetical protein CC1017 [imported] - Caulobacter crescentus

16 55.5 21.5 188 2 E87711 conserved hypothetical protein CC1017 [imported] - Caulobacter crescentus

17 55.5 21.5 288 2 E72671 conserved hypothetical protein CC1017 [imported] - Caulobacter crescentus

18 55.5 21.5 392 2 F84550 conserved hypothetical protein CC1017 [imported] - Caulobacter crescentus

19 55.0 21.3 324 2 A13642 conserved hypothetical protein CC1017 [imported] - Caulobacter crescentus

20 55.0 21.3 377 2 A92655 conserved hypothetical protein CC1017 [imported] - Caulobacter crescentus

21 54.0 20.9 147 2 F83505 conserved hypothetical protein CC1017 [imported] - Caulobacter crescentus

22 54.0 20.9 321 2 H87683 conserved hypothetical protein CC1017 [imported] - Caulobacter crescentus

23 54.0 20.9 1156 2 E63444 conserved hypothetical protein CC1017 [imported] - Caulobacter crescentus

24 53.5 20.7 134 2 A83177 conserved hypothetical protein CC1017 [imported] - Caulobacter crescentus

25 53.5 20.7 352 2 H71294 conserved hypothetical protein CC1017 [imported] - Caulobacter crescentus

26 53.5 20.7 404 2 S63409 conserved hypothetical protein CC1017 [imported] - Caulobacter crescentus

27 53.5 20.7 478 2 C86274 conserved hypothetical protein CC1017 [imported] - Caulobacter crescentus

28 53.0 20.5 210 2 B70399 conserved hypothetical protein CC1017 [imported] - Caulobacter crescentus

29 53.0 20.5 224 2 D83570 conserved hypothetical protein CC1017 [imported] - Caulobacter crescentus

ALIGNMENTS

RESULT 1
E87375 conserved hypothetical protein CC1017 [imported] - Caulobacter crescentus

C.Species: Caulobacter crescentus
C.Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001

C.Accession: E87375 R.Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.

B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon

n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.

Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A.Title: Complete Genome Sequence of Caulobacter crescentus

A.Reference number: A87249; MUID:21173698; PMID:12259647

A.Status: Preliminary

A.Molecule type: DNA

A.Cross-references: GB:AB005673; NID:gi13425041; PIDN:AK25303.1; GSPDB:GN00148

A.Genetics: CCG341

Query Match Score 25.4%; DB 2; Length 528; Pred. No. 2.6; Mismatches 18; Indels 5; Gaps 1;

hypothetical protein plex - Syn
probable 3-dehydro
probable DNA-bind
intermediate filam
hypothetical prote
heat shock protein
NADH-ubiquinone ox
transcription repr
hypothetical prote
hypothetical prote
homeric protein A
probable 2-dehydro
neurofilament-66 -
alpha-internexin -

QY 8 WRGGHVVTEGLAGELBQLRLEHHPOGQREPMMSGCKLG 47
 Db 23 WSAGYVVEITGTYGCPBLNPVALPLAKGPRAPKETACELG 67

RESULT 3

F84015 maltodextrin transport system (maltose/maltodextrin-binding protein) BH2926 [imp] C;Species: *Bacillus halodurans* C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001 R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hira Nucleic Acids Res. 28, 4317-4331, 2000 C;Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and A;Reference number: A83650; MUID:20512582; PMID:11058132 A;Status: preliminary A;Accession: F84015 A;Molecule type: DNA A;Residues: 1-429 <STO> A;Cross-references: GB:AP001517; GB:BA000004; NID:910175500; PIDN:BAB06645.1; GSPDB:GN00154 A;Experimental source: strain C-125 A;Genetics: C;Superfamily: maltose-binding protein

Query Match 23.6%; Score 61; DB 2; Length 429; Best Local Similarity 36.8%; Pred. No. 7.7; Matches 14; Conservative 8; Mismatches 16; Indels 0; Gaps 0; Qy 10 GGTGHWEGLAGELBQLRLEHHPOGQREPMMSGCKLG 47 Db 122 GDTYLGQGLAELBLTPEOLEGYPEGAIDALNYEGAQLG 159

RESULT 4

E64905 probable sugar transport permease protein b1514 - *Escherichia coli* (strain K-12) C;Species: *Escherichia coli* C;Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 01-May-2002 C;Accession: E64905 R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; C;Title: Genome sequence of enterohemorrhagic *Escherichia coli* 0157:H7. A.;Rose, D.J.; Mau, B.; Shao, Y. Science 277, 1455-1462, 1997 A;Reference number: A64720; MUID:97426617; PMID:9278503 A;Status: nucleic acid sequence not shown; translation not shown A;Molecule type: DNA A;Residues: 1-342 <BLAT> A;Cross-references: GB:AE000249; GB:U00096; NID:91787790; PIDN: AAC74587.1; PID:91787793; A;Experimental source: strain K-12, substrain MG1655 C;Superfamily: 1-*arabinose* transport system permease arah C;Keywords: transmembrane protein F;11-27/Domain: transmembrane #status predicted <TM1> F;44-60/Domain: transmembrane #status predicted <TM2> F;65-81/Domain: transmembrane #status predicted <TM3> F;88-104/Domain: transmembrane #status predicted <TM4> F;107-123/Domain: transmembrane #status predicted <TM5> F;155-171/Domain: transmembrane #status predicted <TM6> F;208-224/Domain: transmembrane #status predicted <TM7> F;247-263/Domain: transmembrane #status predicted <TM8> F;266-282/Domain: transmembrane #status predicted <TM9> F;291-307/Domain: transmembrane #status predicted <TM10>

Query Match 23.1%; Score 59.5; DB 2; Length 342; Best Local Similarity 35.0%; Pred. No. 9.3; Matches 14; Conservative 5; Mismatches 10; Indels 11; Gaps 1; Qy 8 WRGGHVVTEGLAGELBQLRLEHHPOGQREPMMSGCKLG 47 Db 135 WTGGKWWIEGLPABLKOUSA-----PLILGVAIG 163

RESULT 5

A;Accession: A90894 Probable transport system permease protein EC82121 [imported] - *Escherichia coli* (strain R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G. DNA Res. 8, 11-22, 2001 C;Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* 0157:H7 and geno A;Reference number: A99229; MUID:2115631; PMID:11250796 A;Accession: A90894 A;Status: preliminary A;Molecule type: DNA A;Residues: 1-342 <HAY> A;Cross-references: GB:BA000007; PIDN: BAB35544.1; PID:913361587; GSPDB:GN00154 A;Experimental source: strain 0157:H7, substrain RIMD 050952 C;Genetics: C;Gene: EC82121 C;Superfamily: 1-*arabinose* transport system permease arah

Query Match 23.1%; Score 59.5; DB 2; Length 342; Best Local Similarity 35.0%; Pred. No. 9.3; Matches 14; Conservative 5; Mismatches 10; Indels 11; Gaps 1; Qy 8 WRGGHVVTEGLAGELBQLRLEHHPOGQREPMMSGCKLG 47 Db 135 WTGGKWWIEGLPABLKOUSA-----PLILGVAIG 163

RESULT 6

H85723 probable transport system permease protein ydeY [imported] - *Escherichia coli* (strain O1 C;Species: *Escherichia coli* C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001 C;Accession: H85723 R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew Miller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca, Nature 409, 529-533, 2001 A;Title: Genome sequence of enterohemorrhagic *Escherichia coli* 0157:H7. A;Reference number: A85880; MUID:21074935; PMID:11200551 A;Accession: H85723 A;Status: preliminary A;Molecule type: DNA A;Residues: 1-342 <STO> A;Cross-references: GB:AB005174; NID:912515152; PIDN: AAG56252.1; GSPDB:GN00145; UNGP:Z21 A;Experimental source: strain O157:H7, substrain EDL333 C;Genetics: C;Gene: ydeY C;Superfamily: 1-*arabinose* transport system permease arah

Query Match 23.1%; Score 59.5; DB 2; Length 342; Best Local Similarity 35.0%; Pred. No. 9.3; Matches 14; Conservative 5; Mismatches 10; Indels 11; Gaps 1; Qy 8 WRGGHVVTEGLAGELBQLRLEHHPOGQREPMMSGCKLG 47 Db 135 WTGGKWWIEGLPABLKOUSA-----PLILGVAIG 163

RESULT 7

T30010 hypothetical protein F58G4.1 - *Caenorhabditis elegans* C;Species: *Caenorhabditis elegans* C;Date: 11-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 08-Sep-2000 C;Accession: T30010 R;Du, Z.; Leibman, D. Submitted to the EMBL Data Library, February 1996 A;Description: The sequence of *C. elegans* cosmid F58G4. A;Reference number: Z20720 A;Accession: T30010 A;Status: preliminary; translated from GB/EMBL/DDBJ

Query Match 22.1%; Score 57; DB 2; Length 663;
 Best Local Similarity 36.1%; Pred. No. 38; Matches 13; Conservative 3; Mismatches 20; Indels 0; Gaps 0;

Qy 1 MDEKTTGGRGHVVEGLAGELBEOLARLEHHPOGQR 36
 Db 154 LDEPLPLAEGGDNLRFARRLSFOQARPATPPQGLR 189

RESULT 12

AI1074 Phosphoglycerate mutase (EC 5.4.2.1) - *Salmonella enterica* subsp. *enterica* serovar *Typhi*
 C;Species: *Salmonella enterica* subsp. *enterica* serovar *Typhi*
 C;Note: this species has also been called *Salmonella typhi*
 C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 27-Oct-2003
 C;Accession: AI1074
 R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, S.; Moulé, S.; O'Gara, P.
 Nature 413, 848-852, 2001.
 A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica* serovar *Th*; T; Compton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, A;Accession: AI1074
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-215 <PAR>
 A;Cross-references: GB:AL513382; PIDN:CAD03416.1; PID:916505684; GSPDB:GN00176
 C;Genetics:
 C;Superfamily: cofactor-dependent phosphoglycerate mutase; phosphoglycerate mutase homologs
 C;Keywords: intramolecular transferase; isomerase

Query Match 21.9%; Score 56.5; DB 2; Length 215;
 Best Local Similarity 26.7%; Pred. No. 14; Matches 16; Conservative 8; Mismatches 23; Indels 13; Gaps 1;

Qy 1 MDEKTTGMR-----GHHVVEGLAGELBEOLARLEHHPOGQREPMSGKLG 47
 Db 96 LDEEEEGGRQLVNGTQDGRIGQGSMQELSDRVHALASCLQPQCSRPLIVSHGIALG 155

RESULT 13

154361 SMX Protein - human
 N;Alternate names: escapes X-chromosome inactivation
 C;Species: Homo sapiens (man)
 C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 17-Mar-2000
 C;Accession: 154361; T38235
 R;Mu, J.; Bilison, J.; Salido, B.; Yen, P.; Mohandas, T.; Shapiro, L.J.
 Hum. Mol. Genet. 3, 153-160, 1994
 A;Title: Isolation and characterization of XE169, a novel human gene that escapes X-inactivation
 A;Reference number: 154361; MUID:94214434; PMID:8162017
 A;Accession: T38235
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: mRNA
 A;Residues: 1-1660 <WNU>
 A;Cross-references: GB:L25270; NID:9457136; PID:9457137
 R;Aguilnik, A.I.; Mitchell, M.J.; Mattei, M.G.; Borsani, G.; Avner, P.A.; Lerner, J.L.; Eumol. Genet. 3, 879-884, 1994
 A;Title: A novel X gene with a widely transcribed Y-linked homologue escapes X-inactivation
 A;Reference number: 138235; MUID:95038739; PMID:7951230
 A;Accession: T38235
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: mRNA
 A;Residues: 280-341, Y', 343-344 <AGU>
 A;Cross-references: EMBL:229650; NID:9472994; PIDN:CAAB2758.1; PID:9472995
 C;Genetics: GDB:SMX; XE169; DDX127E
 A;Cross-references: GDB:9865642; OMIM:314690
 A;Map Position: Xp11.22-Xp11.21
 C;Superfamily: human retinoblastoma binding protein 2

RESULT 14

S43218 hypothetical protein YER126c - yeast (saccharomyces cerevisiae)
 C;Species: Saccharomyces cerevisiae
 C;Date: 31-Dec-1993 #sequence_revision 02-Jun-1994 #text_change 19-Apr-2002
 C;Accession: S43218; S50629
 R;Mulligan, J.T.; Dietrich, F.S.; Hennessey, K.M.; Sehl, P.; Komp, C.; Wei, Y.; Taylor, submitted to the EMBL Data Library, February 1993
 A;Reference number: S30812
 A;Accession: S43218
 A;Molecule type: DNA
 A;Residues: 1-261 <MLT>
 A;Reference number: S50629
 A;Cross-references: GB:U18916; EMBL:L1119; NID:91384128; PIDN: AAC03224.1; PID:9603365
 A;Gene: SGD:KR832
 A;Cross-references: SGD:S0000928
 A;Map position: 5R

Query Match 21.7%; Score 56; DB 2; Length 261;
 Best Local Similarity 26.7%; Pred. No. 19; Matches 12; Conservative 10; Mismatches 19; Indels 4; Gaps 1;

Qy 3 EKTTGGRRGHVVEGLAGELBEOLARLEHHPOGQREPMSGG 43
 Db 40 QKLTGKKGPKFAKKYAKVYSMRKIKAHQSKVKGSSKLDTDG 84

RESULT 15

T36025 conserved hypothetical protein SCC54.12c - *Streptomyces coelicolor*
 C;Species: *Streptomyces coelicolor*
 C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 19-May-2000
 C;Accession: T36025
 R;Seeger, K.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A. submitted to the EMBL Data Library, March 1999
 A;Reference number: 221581
 A;Accession: T36025
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-338 <SEB>
 A;Cross-references: EMBL:AL035591; PIDN: CAB38142.1; GSPDB:GN00070; SCOECD:SCC54.12c
 A;Experimental source: strain A3 (2)
 C;Genetics:
 A;Gene: SCOECD:SCC54.12c
 C;Superfamily: *Bacillus subtilis* conserved hypothetical protein yvcJ

Query Match 21.7%; Score 56; DB 2; Length 338;
 Best Local Similarity 27.0%; Pred. No. 25; Matches 17; Conservative 9; Mismatches 11; Indels 26; Gaps 2;

Qy 11 GHVVEGLAGELB-----QRLARLEHHPOGQREP----MMSGC 44
 Db 166 GRIVGIGIAARELLRLRGADLVIDTSSLNVHLRKAQDAQFAGDQEPLRATWMSFGF 225

Qy 45 KLG 47

Fri Jul 16 08:31:15 2004

us-10-004-381-25.rpr

Page 5

Db 226 KYG 228

Search completed: July 15, 2004, 20:36:20
Job time : 13.7176 secs

This Page Blank (uspto)

Copyright (c) 1993 - 2004 Compugen Ltd.

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: July 15, 2004, 20:35:51 ; Search time 44.2353 Seconds (without alignments)

332.096 Million cell updates/sec

Title: US-10-004-381-25

Perfect score: 258

Sequence: 1 MDEKTTGWRGGHVVVEGLAGE.....LEHHPQGOREPMMGGCKLG 47

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1285345 seqs, 31256033 residues

Total number of hits satisfying chosen parameters: 1285345

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

database : Published Applications AA:*

1: /cgn2_6/ptodata/1/pubpaal/US07_PUBCOMB.pep:*

2: /cgn2_6/ptodata/1/pubpaal/PCN_NEW_PUB_COMB.pep:*

3: /cgn2_6/ptodata/1/pubpaal/US06_NEW_PUB_COMB.pep:*

4: /cgn2_6/ptodata/1/pubpaal/US06_PUBCOMB.pep:*

5: /cgn2_6/ptodata/1/pubpaal/US07_NEW_PUB_COMB.pep:*

6: /cgn2_6/ptodata/1/pubpaal/PCUTUS_PUBCOMB.pep:*

7: /cgn2_6/ptodata/1/pubpaal/US08_NEW_PUB_COMB.pep:*

8: /cgn2_6/ptodata/1/pubpaal/US08_PUBCOMB.pep:*

9: /cgn2_6/ptodata/1/pubpaal/US09_PUBCOMB.pep:*

10: /cgn2_6/ptodata/1/pubpaal/US09_PUBCOMB.pep:*

11: /cgn2_6/ptodata/1/pubpaal/US09C_PUBCOMB.pep:*

12: /cgn2_6/ptodata/1/pubpaal/US09_NEW_PUB_COMB.pep:*

13: /cgn2_6/ptodata/1/pubpaal/US10_PUBCOMB.pep:*

14: /cgn2_6/ptodata/1/pubpaal/US10B_PUBCOMB.pep:*

15: /cgn2_6/ptodata/1/pubpaal/US10C_PUBCOMB.pep:*

16: /cgn2_6/ptodata/1/pubpaal/US10C_PUBCOMB.pep:*

17: /cgn2_6/ptodata/1/pubpaal/US60_NEW_PUB_COMB.pep:*

18: /cgn2_6/ptodata/1/pubpaal/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	258	100.0	47	13	US-10-004-381-25
2	243	94.2	47	13	US-10-004-381-25
3	241	93.4	61	13	US-10-004-381-24
4	234	90.7	75	13	US-10-004-381-23
5	227	88.0	89	13	US-10-004-381-22
6	214	82.9	101	13	US-10-004-381-21
7	211	81.8	101	13	US-10-004-381-19
8	208	80.6	38	13	US-10-004-381-35
9	208	80.6	426	13	US-10-004-381-41
10	208	80.6	479	13	US-10-004-381-38
11	199	77.1	39	13	US-10-004-381-27
12	152	58.9	29	13	US-10-004-381-29
13	133	51.6	89	13	US-10-004-381-28
14	36.4	44	13	US-10-004-381-34	
15	27.1	102	13	US-10-004-381-11	

ALIGNMENTS

RESULT 1

US-10-004-381-25

; Sequence 25, Application US/10004381

; Publication No. US20020155578A1

; GENERAL INFORMATION:

; APPLICANT: SZOSTAK, JACK W

; APPLICANT: WILSON, DAVID S.

; APPLICANT: KEFFE, ANTHONY D.

; TITLE OF INVENTION: STREPTAVIDIN-BINDING PEPTIDES AND USES

; FILE REFERENCE: 00786/3/002

; CURRENT APPLICATION NUMBER: US/10/004, 381

; CURRENT FILING DATE: 2001-10-31

; PRIORITY APPLICATION NUMBER: US 60/244,541

; PRIORITY FILING DATE: 2000-10-31

; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO: 25

; LENGTH: 47

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURES:

; OTHER INFORMATION: selected peptide

US-10-004-381-25

Query Match 100.0%; Score 258; DB 13; Length 47;

Best Local Similarity 100.0%; Pred. No. 7e-26;

Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MDEKTTGWRGGHVVVEGLAGE.....LEHHPQGOREPMMGGCKLG 47

Db 1 MDEKTTGWRGGHVVVEGLAGE.....LEHHPQGOREPMMGGCKLG 47

RESULT 2

US-10-004-381-26

; Sequence 26, Application US/10004381

; Publication No. US20020155578A1

GENERAL INFORMATION:
; APPLICANT: SZOSTAK, JACK W.
; APPLICANT: WILSON, DAVID S.
; APPLICANT: KEEFE, ANTHONY D.
; TITLE OF INVENTION: STREPTAVIDIN-BINDING PEPTIDES AND USES
; FILE REFERENCE: 00786/388002
; CURRENT APPLICATION NUMBER: US/10/004,381
; CURRENT FILING DATE: 2001-10-31
; PRIOR APPLICATION NUMBER: US 60/244,541
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO: 26
; LENGTH: 47
; TYPE: PRT
; FEATURE: OTHER INFORMATION: Artificial Sequence
; OTHER INFORMATION: selected peptide
; US-10-004-381-26

Query Match 94.2%; Score 243; DB 13; Length 47;
Best Local Similarity 95.7%; Pred. No. 6e-24;
Matches 45; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MDEKTTGWRGHHVVEGLAGELQQLARLEHHPQOREPMMGGCKLG 47
Db 1 MDEKTTGWRGHHVVEGLAGELQQLARLEHHPQOREPMMGGCKLG 47

RESULT 3
US-10-004-381-24
; Sequence 24, Application US/10004381
; Publication No. US2002015578A1
; GENERAL INFORMATION:
; APPLICANT: SZOSTAK, JACK W.
; APPLICANT: WILSON, DAVID S.
; APPLICANT: KEEFE, ANTHONY D.
; TITLE OF INVENTION: STREPTAVIDIN-BINDING PEPTIDES AND USES
; FILE REFERENCE: 00786/388002
; CURRENT APPLICATION NUMBER: US/10/004,381
; CURRENT FILING DATE: 2001-10-31
; PRIOR APPLICATION NUMBER: US 60/244,541
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO: 24
; LENGTH: 61
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE: OTHER INFORMATION: selected peptide
; OTHER INFORMATION: selected peptide
; US-10-004-381-24

Query Match 93.4%; Score 241; DB 13; Length 61;
Best Local Similarity 77.0%; Pred. No. 1.5e-23;
Matches 47; Conservative 0; Mismatches 0; Indels 14; Gaps 1;

Qy 1 MDEKTTGWRGHHVVEGLAGELQQLARLEHHPQOREPMMGGCKLG 46
Db 1 MDEKTTGWRGHHVVEGLAGELQQLARLEHHPQOREPMMGGCKLG 60

RESULT 4
US-10-004-381-23
; Sequence 23, Application US/10004381
; Publication No. US2002015578A1
; GENERAL INFORMATION:

GENERAL INFORMATION:
; APPLICANT: SZOSTAK, JACK W.
; APPLICANT: WILSON, DAVID S.
; APPLICANT: KEEFE, ANTHONY D.
; TITLE OF INVENTION: STREPTAVIDIN-BINDING PEPTIDES AND USES
; FILE REFERENCE: 00786/388002
; CURRENT APPLICATION NUMBER: US/10/004,381
; CURRENT FILING DATE: 2001-10-31
; PRIOR APPLICATION NUMBER: US 60/244,541
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO: 23
; LENGTH: 75
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE: OTHER INFORMATION: selected peptide
; OTHER INFORMATION: selected peptide
; US-10-004-381-23

Query Match 90.7%; Score 234; DB 13; Length 75;
Best Local Similarity 62.7%; Pred. No. 1.5e-22;
Matches 47; Conservative 0; Mismatches 0; Indels 28; Gaps 1;

Qy 1 MDEKTTGWRGHHVVEGLAGELQQLARLEHHPQOREPMMGGCKLG 38
Db 1 MDEKTTGWRGHHVVEGLAGELQQLARLEHHPQOREPMMGGCKLG 60

RESULT 5
US-10-004-381-22
; Sequence 22, Application US/10004381
; Publication No. US2002015578A1
; GENERAL INFORMATION:
; APPLICANT: SZOSTAK, JACK W.
; APPLICANT: WILSON, DAVID S.
; APPLICANT: KEEFE, ANTHONY D.
; TITLE OF INVENTION: STREPTAVIDIN-BINDING PEPTIDES AND USES
; FILE REFERENCE: 00786/388002
; CURRENT APPLICATION NUMBER: US/10/004,381
; CURRENT FILING DATE: 2001-10-31
; PRIOR APPLICATION NUMBER: US 60/244,541
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO: 22
; LENGTH: 89
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE: OTHER INFORMATION: selected peptide
; OTHER INFORMATION: selected peptide
; US-10-004-381-22

Query Match 88.0%; Score 227; DB 13; Length 89;
Best Local Similarity 52.8%; Pred. No. 1.4e-21;
Matches 47; Conservative 0; Mismatches 0; Indels 42; Gaps 1;

Qy 1 MDEKTTGWRGHHVVEGLAGELQQLARLEHHPQOREPMMGGCKLG 38
Db 1 MDEKTTGWRGHHVVEGLAGELQQLARLEHHPQOREPMMGGCKLG 60

RESULT 6
US-10-004-381-21

Sequence 21, Application US/10004381
 Publication No. US20020155578A1
 GENERAL INFORMATION:
 APPLICANT: ZOSTAK, JACK W.
 APPLICANT: WILSON, DAVID S.
 APPLICANT: KEEFE, ANTHONY D.
 TITLE OF INVENTION: STREPTAVIDIN-BINDING PEPTIDES AND USES
 TITLE OF INVENTION: THEREOF
 FILE REFERENCE: 00786/388002
 CURRENT APPLICATION NUMBER: US/10/004,381
 CURRENT FILING DATE: 2001-10-31
 PRIORITY APPLICATION NUMBER: US 60/244,541
 PRIORITY FILING DATE: 2000-10-31
 NUMBER OF SEQ ID NOS: 41
 SOFTWARE: FastSEQ for Windows Version 4.0
 SEQ ID NO: 21
 LENGTH: 101
 TYPE: PRT
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: selected peptide
 US-10-004-381-21

Query Match 82.9%; Score 214; DB 13; Length 101;
 Best Local Similarity 44.6%; Pred. No. 7.8e-20;
 Matches 45; Conservative 2; Mismatches 0; Indels 54; Gaps 1;
 APPLICANT: SZOSTAK, JACK W.
 APPLICANT: WILSON, DAVID S.
 APPLICANT: KEEFE, ANTHONY D.
 TITLE OF INVENTION: STREPTAVIDIN-BINDING PEPTIDES AND USES
 FILE REFERENCE: 00786/388002
 CURRENT FILING DATE: 2001-10-31
 PRIORITY APPLICATION NUMBER: US 60/244,541
 NUMBER OF SEQ ID NOS: 41
 SOFTWARE: FastSEQ for Windows Version 4.0
 SEQ ID NO: 19
 LENGTH: 101
 TYPE: PRT
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: selected peptide
 US-10-004-381-19

Sequence 19, Application US/10004381
 Publication No. US20020155578A1
 GENERAL INFORMATION:
 APPLICANT: ZOSTAK, JACK W.
 APPLICANT: WILSON, DAVID S.
 APPLICANT: KEEFE, ANTHONY D.
 TITLE OF INVENTION: STREPTAVIDIN-BINDING PEPTIDES AND USES
 FILE REFERENCE: 00786/388002
 CURRENT APPLICATION NUMBER: US/10/004,381
 CURRENT FILING DATE: 2001-10-31
 PRIORITY APPLICATION NUMBER: US 60/244,541
 NUMBER OF SEQ ID NOS: 41
 SOFTWARE: FastSEQ for Windows Version 4.0
 SEQ ID NO: 19
 LENGTH: 101
 TYPE: PRT
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: selected peptide
 US-10-004-381-19

Query Match 81.8%; Score 211; DB 13; Length 101;
 Best Local Similarity 95.0%; Pred. No. 1.9e-19;
 Matches 38; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 APPLICANT: SZOSTAK, JACK W.
 APPLICANT: WILSON, DAVID S.
 APPLICANT: KEEFE, ANTHONY D.
 TITLE OF INVENTION: THEREOF
 FILE REFERENCE: 00786/388002
 CURRENT APPLICATION NUMBER: US/10/004,381
 CURRENT FILING DATE: 2001-10-31
 PRIORITY APPLICATION NUMBER: US 60/244,541
 NUMBER OF SEQ ID NOS: 40
 SOFTWARE: FastSEQ for Windows Version 4.0
 SEQ ID NO: 19
 LENGTH: 101
 TYPE: PRT
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: selected peptide
 US-10-004-381-19

Query Match 80.6%; Score 208; DB 13; Length 38;
 Best Local Similarity 100.0%; Pred. No. 2.3e-18;
 Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 APPLICANT: SZOSTAK, JACK W.
 APPLICANT: WILSON, DAVID S.
 APPLICANT: KEEFE, ANTHONY D.
 TITLE OF INVENTION: STREPTAVIDIN-BINDING PEPTIDES AND USES
 FILE REFERENCE: 00786/388002
 CURRENT APPLICATION NUMBER: US/10/004,381
 CURRENT FILING DATE: 2001-10-31
 PRIORITY APPLICATION NUMBER: US 60/244,541
 NUMBER OF SEQ ID NOS: 41
 SOFTWARE: FastSEQ for Windows Version 4.0
 SEQ ID NO: 41
 LENGTH: 426
 TYPE: PRT
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: designed protein
 US-10-004-381-41

Query Match 80.6%; Score 208; DB 13; Length 426;
 Best Local Similarity 100.0%; Pred. No. 2.3e-18;
 Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 APPLICANT: SZOSTAK, JACK W.
 APPLICANT: WILSON, DAVID S.
 APPLICANT: KEEFE, ANTHONY D.
 TITLE OF INVENTION: THEREOF
 FILE REFERENCE: 00786/388002
 CURRENT APPLICATION NUMBER: US/10/004,381
 CURRENT FILING DATE: 2001-10-31
 PRIORITY APPLICATION NUMBER: US 60/244,541
 NUMBER OF SEQ ID NOS: 40
 SOFTWARE: FastSEQ for Windows Version 4.0
 SEQ ID NO: 40
 LENGTH: 426
 TYPE: PRT
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: designed protein
 US-10-004-381-40

Query Match 80.6%; Score 208; DB 13; Length 426;
 Best Local Similarity 100.0%; Pred. No. 2.3e-18;
 Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 APPLICANT: SZOSTAK, JACK W.
 APPLICANT: WILSON, DAVID S.
 APPLICANT: KEEFE, ANTHONY D.
 TITLE OF INVENTION: STREPTAVIDIN-BINDING PEPTIDES AND USES
 FILE REFERENCE: 00786/388002
 CURRENT APPLICATION NUMBER: US/10/004,381
 CURRENT FILING DATE: 2001-10-31
 PRIORITY APPLICATION NUMBER: US 60/244,541
 NUMBER OF SEQ ID NOS: 41
 SOFTWARE: FastSEQ for Windows Version 4.0
 SEQ ID NO: 35
 LENGTH: 38
 TYPE: PRT
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: selected peptide
 US-10-004-381-35

Query Match 80.6%; Score 208; DB 13; Length 38;
 Best Local Similarity 100.0%; Pred. No. 1.6e-19;
 Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 APPLICANT: SZOSTAK, JACK W.
 APPLICANT: WILSON, DAVID S.
 APPLICANT: KEEFE, ANTHONY D.
 TITLE OF INVENTION: STREPTAVIDIN-BINDING PEPTIDES AND USES
 FILE REFERENCE: 00786/388002
 CURRENT APPLICATION NUMBER: US/10/004,381
 CURRENT FILING DATE: 2001-10-31
 PRIORITY APPLICATION NUMBER: US 60/244,541
 NUMBER OF SEQ ID NOS: 41
 SOFTWARE: FastSEQ for Windows Version 4.0
 SEQ ID NO: 38
 LENGTH: 38
 TYPE: PRT
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: selected peptide
 US-10-004-381-38

; TITLE OF INVENTION: STREPTAVIDIN-BINDING PEPTIDES AND USES
 ; TITLE OF INVENTION: THEREOF
 ; FILE REFERENCE: 00786/388002
 ; CURRENT APPLICATION NUMBER: US/10/004, 381
 ; CURRENT FILING DATE: 2001-10-31
 ; PRIORITY APPLICATION NUMBER: US 60/244, 541
 ; PRIORITY FILING DATE: 2000-10-31
 ; NUMBER OF SEQ ID NOS: 41
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO 38
 ; LENGTH: 479
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Designed protein
 ; US-10-004-381-38

Query Match 80.6%; Score 208; DB 13; Length 479;
 Best Local Similarity 100.0%; Pred. No. 2.6e-18; Mismatches 0; Indels 0; Gaps 0;
 Matches 38; Conservative

QY 1 MDRKTTGGRGHRGVEGLAGLEBOLARLEHHPOGOREP 38
 Db 378 MDRKTTGGRGHRGVEGLAGLEBOLARLEHHPOGOREP 415

RESULT 11

US-10-004-381-27
 ; Sequence 27, Application US/10004381
 ; Publication No. US20020155578A1

; GENERAL INFORMATION:
 ; APPLICANT: SZOSTAK, JACK W.
 ; APPLICANT: WILSON, DAVID S.
 ; APPLICANT: KEEFE, ANTHONY D.
 ; TITLE OF INVENTION: STREPTAVIDIN-BINDING PEPTIDES AND USES
 ; TITLE OF INVENTION: THEREOF
 ; FILE REFERENCE: 00786/388002
 ; CURRENT APPLICATION NUMBER: US/10/004, 381
 ; CURRENT FILING DATE: 2001-10-31
 ; PRIORITY APPLICATION NUMBER: US 60/244, 541
 ; PRIORITY FILING DATE: 2000-10-31
 ; NUMBER OF SEQ ID NOS: 41
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO 27
 ; LENGTH: 39

; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; OTHER INFORMATION: selected peptide
 ; US-10-004-381-27

Query Match 77.1%; Score 199; DB 13; Length 39;
 Best Local Similarity 100.0%; Pred. No. 2.4e-18; Mismatches 0; Indels 0; Gaps 0;
 Matches 37; Conservative

QY 11 GRVVEGLAGLEBOLARLEHHPOGOREPMMGGCKLG 47
 Db 3 GHVVEGLAGLEBOLARLEHHPOGOREPMMGGCKLG 39

RESULT 12

US-10-004-381-29
 ; Sequence 29, Application US/10004381
 ; Publication No. US20020155578A1

; GENERAL INFORMATION:
 ; APPLICANT: SZOSTAK, JACK W.
 ; APPLICANT: WILSON, DAVID S.
 ; APPLICANT: KEEFE, ANTHONY D.
 ; TITLE OF INVENTION: STREPTAVIDIN-BINDING PEPTIDES AND USES
 ; FILE REFERENCE: 00786/388002
 ; CURRENT APPLICATION NUMBER: US/10/004, 381
 ; CURRENT FILING DATE: 2001-10-31

Query Match 80.6%; Score 208; DB 13; Length 479;
 Best Local Similarity 100.0%; Pred. No. 2.6e-18; Mismatches 0; Indels 0; Gaps 0;
 Matches 38; Conservative

QY 1 MDRKTTGGRGHRGVEGLAGLEBOLARLEHHPOGOREP 38
 Db 378 MDRKTTGGRGHRGVEGLAGLEBOLARLEHHPOGOREP 415

RESULT 13

US-10-004-381-28
 ; Sequence 28, Application US/10004381
 ; Publication No. US20020155578A1

; GENERAL INFORMATION:
 ; APPLICANT: SZOSTAK, JACK W.
 ; APPLICANT: WILSON, DAVID S.
 ; APPLICANT: KEEFE, ANTHONY D.
 ; TITLE OF INVENTION: STREPTAVIDIN-BINDING PEPTIDES AND USES
 ; TITLE OF INVENTION: THEREOF
 ; FILE REFERENCE: 00786/388002
 ; CURRENT APPLICATION NUMBER: US/10/004, 381
 ; CURRENT FILING DATE: 2001-10-31
 ; PRIORITY APPLICATION NUMBER: US 60/244, 541
 ; PRIORITY FILING DATE: 2000-10-31
 ; NUMBER OF SEQ ID NOS: 41
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO 28
 ; LENGTH: 89

; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: selected peptide
 ; US-10-004-381-28

Query Match 51.6%; Score 133; DB 13; Length 89;
 Best Local Similarity 35.6%; Pred. No. 2e-09; Mismatches 2; Indels 54; Gaps 1;
 Matches 31; Conservative

QY 15 EGLACBEBOLARLEHHPOGOREPMMGGCKLG 47
 Db 3 EGLACBEBOLARLEHHPOGOREPMMGGCKLG 62

RESULT 14

US-10-004-381-34
 ; Sequence 34, Application US/10004381
 ; Publication No. US20020155578A1

; GENERAL INFORMATION:
 ; APPLICANT: SZOSTAK, JACK W.
 ; APPLICANT: WILSON, DAVID S.
 ; APPLICANT: KEEFE, ANTHONY D.
 ; TITLE OF INVENTION: STREPTAVIDIN-BINDING PEPTIDES AND USES
 ; FILE REFERENCE: 00786/388002
 ; CURRENT APPLICATION NUMBER: US/10/004, 381
 ; CURRENT FILING DATE: 2001-10-31
 ; PRIORITY APPLICATION NUMBER: US 60/244, 541

; PRIORITY FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 34
; LENGTH: 44
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: designed peptide
US-10-004-381-34

RESULT 15
US-10-004-381-11
; Sequence 11, Application US/10004381
; GENERAL INFORMATION:
; Publication No. US2002015578A1
; APPLICANT: SZOSTAK, JACK W.
; APPLICANT: WILSON, DAVID S.
; APPLICANT: KEEFE, ANTHONY D.
; TITLE OF INVENTION: STREETAVIDIN-BINDING PEPTIDES AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 00786/38802
; CURRENT APPLICATION NUMBER: US/10/004,381
; PRIOR APPLICATION NUMBER: US 60/244,541
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 102
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: selected peptide
US-10-004-381-11

Query Match 27.1%; Score 70; DB 13; Length 102;
Best Local Similarity 42.9%; Pred. No. 0.33; Mismatches 6; Indels 14; Gaps 3;
Matches 18; Conservative 4; Mismatches 6; Indels 14; Gaps 3;

QY 1 MDEKTTGWI-RGSHVVEGLAEQLRARLHHHQGRBPPMM 40
Db 1 MDEK-THHWRVHLADG-----LEQHQGQRRLV 30

Search completed: July 15, 2004, 20:45:03
Job time : 45.2353 secs

This Page Blank (up to)

RESULT 4
US-09-134-000C-5025
; Sequence 5025, Application US/09134000C
; Patent No. 617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778

Query Match 24.0%; Score 62; DB 4; Length 356;
Best Local Similarity 41.0%; Pred. No. 1,8; Mismatches 5; Indels 6; Gaps 2;
Matches 16; Conservative 5; Mismatches 12; Indels 6; Gaps 2;

Qy 7 GMRGGHVGEGAGLEQLRARLHHPQ-GQREPMMSGG 43
Db 183 GHRGGHAGAGEPQ---RATBHQHPEAVGQRQQTAEG 217

RESULT 5
US-09-252-991A-18469
; Sequence 18469, Application US/09252991A
; Patent No. 655175
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196-136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 18469
; LENGTH: 356
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-19549

Query Match 24.4%; Score 63; DB 3; Length 544;
Best Local Similarity 35.1%; Pred. No. 2,1; Mismatches 6; Indels 8; Gaps 1;
Matches 13; Conservative 6; Mismatches 10; Indels 8; Gaps 1;

Qy 8 WRGGHVVVEG-----LAGELQQLRARLHHPQGQR 36
Db 236 WLSGEVLEGLERWVTTKUSGALLRARRVVPVDHSPAGRR 272

RESULT 3
US-09-252-991A-19549
; Sequence 19549, Application US/09252991A
; Patent No. 655175
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196-136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 18469
; LENGTH: 356
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-18469

Query Match 23.6%; Score 61; DB 4; Length 239;
Best Local Similarity 48.5%; Pred. No. 1,5; Mismatches 3; Indels 4; Gaps 2;
Matches 16; Conservative 3; Mismatches 10; Indels 4; Gaps 2;

Qy 8 WRGGHVVVEGAGLQLRARLHHPQGQRPM 40
Db 34 W-GGEVITGLMASMEOERAKEHE---GREMM 62

RESULT 6
US-09-252-991A-21794
; Sequence 21794, Application US/09252991A
; Patent No. 655175
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196-136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 21794
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (405)
; OTHER INFORMATION: Identity of amino acid at the above locations are unknown.

RESULT 11
 US-09-252-991A-27885 Application US/09252991A
 ; Sequence 27885, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: MARC J. RUBENFIELD et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIORITY FILING DATE: 1998-02-18
 ; PRIORITY APPLICATION NUMBER: US 60/094,190
 ; PRIORITY FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO: 27885
 ; LENGTH: 286
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 ; US-09-252-991A-27885

Query Match 22.5%; Score 58; DB 4; Length 286;
 Best Local Similarity 33.3%; Pred. No. 4.8; Mismatches 15; Conservative 7; Indels 0; Gaps 0;
 Matches 23; General Information:
 Qy 3 EKTRGWRSGHUVVSGLAGELFOLRARLEHHRPGOREPBMMSGCKLG 47
 Db 78 EHTHDPRGESERVEAGGFVRQQQGLVLVDDGAGNAIDLILLAGKVG 122

RESULT 12
 US-09-252-991A-22131 Application US/09252991A
 ; Sequence 22131, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: MARC J. RUBENFIELD et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; PRIORITY FILING DATE: 1999-02-18
 ; PRIORITY APPLICATION NUMBER: US 60/094,190
 ; PRIORITY FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO: 22131
 ; LENGTH: 628
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 ; US-09-252-991A-22131

Query Match 22.5%; Score 58; DB 4; Length 628;
 Best Local Similarity 44.4%; Pred. No. 12; Mismatches 16; Conservative 1; Indels 8; Gaps 1;
 Matches 11; General Information:
 Qy 9 RGGHVE-----GLAGELFOLRARLEHHRPGOREPBMMSGCKLG 36
 Db 126 RGGHAAVEVGGLAVGGLAVELGAVGPGADHPVGFR 161

RESULT 13
 US-09-412-545-2 Application US/09412545
 ; Sequence 2, Application US/09412545
 ; Patent No. 625095
 ; GENERAL INFORMATION:
 ; APPLICANT: PRESCOTT, Stephen M.
 ; APPLICANT: Ding, Li

RESULT 14
 US-09-252-991A-30737 Application US/09252991A
 ; Sequence 30737, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: MARC J. RUBENFIELD et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIORITY FILING DATE: 1998-02-18
 ; PRIORITY APPLICATION NUMBER: US 60/094,190
 ; PRIORITY FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO: 30737
 ; LENGTH: 296
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 ; US-09-252-991A-30737

Query Match 22.1%; Score 57; DB 4; Length 296;
 Best Local Similarity 46.2%; Pred. No. 6.9; Mismatches 12; Conservative 2; Indels 0; Gaps 0;
 Matches 12; General Information:
 Qy 18 AGELFOLRARLEHHRPGOREPBMMSGCKLG 43
 Db 55 AGVLDLHLRVAEEHHPRMLRCQQFQVGG 80

RESULT 15
 US-08-311-731A-3 Application US/08311731A
 ; Sequence 3, Application US/08311731A
 ; Patent No. 6583266
 ; GENERAL INFORMATION:
 ; APPLICANT: SMITH, DOUGLAS
 ; APPLICANT: MAO, JEN-I
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
 ; TITLE OF INVENTION: RELATING TO MYCOBACTERIUM TUBERCULOSIS AND LAPRAE FOR
 ; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
 ; NUMBER OF SEQUENCES: 411
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSE: WOLF, GREENFIELD & SACKS, P.C.
 ; STREET: 600 ATLANTIC AVENUE
 ; CITY: BOSTON
 ; STATE: MASSACHUSETTS
 ; COUNTRY: USA
 ; ZIP: 02210

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/311,731A
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: GATES, EDWARD R.
REGISTRATION NUMBER: 31,616
REFERENCE/DOCKET NUMBER: C0044/7125
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/720-3500
TELEFAX: 617/720-2441
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 495 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: Protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: MYCOBACTERIUM TUBERCULOSIS
US-08-311-731A-3

Query Match 22.1%; Score 57; DB 4; Length 495;
Best Local Similarity 36.1%; Pred. No. 12;
Matches 13; Conservative 9; Mismatches 12; Indels 2; Gaps 1;

Qy 12 HVVEGLAGELQELQRARLREHHPOCQREP--MMSGCK 45
Db 22 HVVKVAVELDEIRARLREYEQARATEPVAVVGCR 57

Search completed: July 15, 2004, 20:37:00
Job time : 17.0353 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using SW model

Run on: July 15, 2004, 20:30:55 ; Search time 29.5059 Seconds
(without alignments)
406.349 Million cell updates/sec

Title: US-10-004-381-25_COPY_1_38

Perfect score: 208

Sequence: 1 MDEKTTGWRGGHVVVEGLAGLEQLRARLHHPOQREP 38

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : SPTREMBL_25;*

- 1: sp_archea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp Rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriop:*
- 17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match Length	DB ID	Description	RESULT 1	ALIGNMENTS
1	66.5	32.0	739	16 Q82BY9	082BY9 STREPTOMYCE	082BY9 PRELIMINARY; ID: 082BY9; AC: 082BY9; DT: 01-JUN-2003 (TREMBLrel: 24, last sequence update) DR: 01-JUN-2003 (TREMBLrel: 24, last annotation update) DE: putative regulatory protein. GN: SAV5555.	
2	65.5	31.5	189	2 Q8KR07	Q8KR07 rhodococcus		
3	61.	29.3	773	2 033541	033541 endosymbiont		
4	60.	28.8	588	16 Q8DH85	Q8DH85 cyanobacter		
5	59.	28.4	139	4 Q96MD1	Q96MD1 homo sapien		
6	59.	28.4	342	16 Q8XAY8	Q8XAY8 escherichia		
7	59.	28.4	342	16 Q83RDB	Q83RDB shigella fl		
8	59.	28.4	1083	3 Q94189	Q94189 cryptococcus		
9	59.	28.4	2360	3 Q94188	Q94188 cryptococcus		
10	58.5	28.1	274	16 Q8P4G4	Q8P4G4 xanthomonas		
11	58.	27.9	333	4 Q8IU04	Q8IU04 homo sapien		
12	58.	27.9	491	2 Q92GD9	Q92GD9 streptomyces		
13	58.	27.9	1161	16 Q910L0	Q910L0 streptomyces		
14	58.	27.9	1181	16 Q82DQ5	Q82DQ5 streptomyces		
15	57.	27.4	229	16 Q7U554	Q7U554 synchococcus		
16	57.	27.4	489	16 Q8FLJ7	Q8FLJ7 corynebacter		

DR SMART; SM00331; PP2C_SIG; 1.
 KW Complete proteome.
 SQ 739 AA; 77478 MW; 638E3F93EC67A2B6 CRC64;

Query Match Similarity 32.0%; Score 66.5; DB 16; Length 739;
 Best Local Similarity 32.6%; Pred. No. 7, 5; Mismatches 13; Indels 11; Gaps 1;

Matches 14; Conservative 5; MisMatches 13; Indels 11; Gaps 1;

Qy 1 MDEKTTGMRG-----GHVVEGLAGELQQLRLEHHIP 32
 DB 341 LDEBEATGWRGPERGMARGLRPLRKYRWWHGSENTRIEDLRRALERHP 383

RESULT 2
 Q8KRD7 PRELIMINARY; PRT; 189 AA.
 ID Q8KRD7
 AC Q8KRD7; 2002 (TREMBrel. 22, Created)
 DT 01-OCT-2002 (TREMBrel. 22, Last sequence update)
 DT 01-MAR-2003 (TREMBrel. 23, Last annotation update)
 DE Transcription repressor, KF-α-like protein.
 OS Rhodococcus erythropolis.

OC Bacteria; Actinobacteria; Actinomycetales;
 OC Corynebacteriaceae; Nocardiaceae; Rhodococcus.
 OX NCBI_TaxID=1833;
 RN [1]

SEQUENCE FROM N.A.

RC STRAIN=MPS0;
 RX MEDLINE=22033458; PubMed=12089004;
 RA Trott S., Burger S., Calaminus C., Stolz A.;
 RT "Cloning and Heterologous Expression of an Enantioselective Amidase
 from Rhodococcus erythropolis Strain MPS0";
 RL Appl. Environ. Microbiol. 68:3279-3286 (2002).

DR EMBL; AY026386; AACID1725.1; 189 AA; 20101 MW; 4B8B5D00A093FFD CRC64;
 SQ SEQUENCE

Query Match 31.5%; Score 65.5; DB 2; Length 189;
 Best Local Similarity 45.5%; Pred. No. 2, 2;
 Matches 20; Conservative 2; Mismatches 13; Indels 9; Gaps 1;

Qy 3 EKTTGWRGHHV-----VEGLAGELQQLRLEHHIPQORE 37
 DB 104 EQATAERDGAIVDARDALGEVEGLRVELEQKQQLTARDQARE 147

RESULT 3
 Q33541 PRELIMINARY; PRT; 773 AA.
 ID Q33541
 AC 033541;
 DT 01-JAN-1998 (TREMBrel. 05, Created)
 DT 01-JAN-1998 (TREMBrel. 05, Last sequence update)
 DT 01-OCT-2003 (TREMBrel. 25, Last annotation update)
 DE Sensor protein ResA.
 RN RSSA.

OS endosymbiont of *Riftia pachyptila*.
 OC Bacteria; Proteobacteria; Gammaproteobacteria;
 OC sulfur-oxidizing symbionts.

NCBI_TaxID=54396;
 RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=97438512; PubMed=2293000;
 RA Hughes D.S., Helbeck H., Stein J.L.;
 RT "A histidine protein kinase homolog from the endosymbiont of the
 hydrothermal vent tubeworm *Riftia pachyptila*";
 RL Appl. Environ. Microbiol. 63:3949-3958 (1997).

CC KINASES.
 DR EMBL; U93704; AAB71131.1; -.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:005254; F:ATP binding; IEA.
 DR GO; GO:03577; F:DNA binding; IEA.
 DR GO; GO:0016301; F:kinase activity; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.

DR SMART; SM00155; F:two-component response regulator activity; IEA.
 DR GO; GO:0000155; F:two-component sensor molecule activity; IEA.
 DR GO; GO:000760; P:sensory perception; IEA.
 DR GO; GO:0000160; P:two-component signal transduction system (p. . .); IEA.

DR InterPro; IPR003594; Atbind AtPase.
 DR InterPro; IPR004358; Bact_Sens_Pr_-.
 DR InterPro; IPR005467; His_Kinase.
 DR InterPro; IPR003661; His_KinA_N.
 DR InterPro; IPR001789; Response_reg.
 DR Pfam; PF02518; Hatase_C; 1.
 DR Pfam; PF00512; HisKA; 1.
 DR Pfam; PF00072; HisKA; 1.
 DR PRINTS; PRO0034; BCURLENSOR.
 DR PRODOM; PRO00019; Response_reg; 1.
 DR SMART; SM0037; HAMPase_C; 1.
 DR SMART; SM00388; HisKA; 1.
 DR PROSITE; PS50109; HIS_KIN; 1.
 DR PROSITE; PS50110; RESPONSE_REGULATORY; 1.
 DR Kinase; Phosphorylation; Sensory transduction; Transferase.
 KW SEQUENCE 733 AA; 87086 MW; 2C5643B64F373B27 CRC64;

Query Match 29.3%; Score 61; DB 2; Length 773;
 Best Local Similarity 43.3%; Pred. No. 39; Mismatches 13; Conservative 2; Indels 15; Gaps 0; Gaps 0;

Qy 8 WRGGHVEGLAGELQQLRLEHHIPQORE 37
 DB 346 WRGPBRIEELGEKLTAALAEHTAHHTQAQRE 375

RESULT 4
 Q8DH85 PRELIMINARY; PRT; 588 AA.
 ID Q8DH85
 AC Q8DH85;
 DT 01-MAR-2003 (TREMBrel. 23, Created)
 DT 01-OCT-2003 (TREMBrel. 25, Last sequence update)
 DR Penicillin-binding protein.
 RN TUR2074.

OS Synechococcus elongatus (Thermosynechococcus elongatus).
 OC Bacteria; Cyanobacteria; Chroococcales; Synechococcus.
 OX NCBI_TaxID=32046;
 RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=BP-1;
 RX MEDLINE=2222544; PubMed=12240834;
 RA Nakamura Y., Kaneko T., Sato S., Ikeuchi M., Katoh H., Sasamoto S.,
 RA Watanabe A., Iriguchi M., Kawashima K., Kimura T., Kishida Y.,
 RA Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Nakazaki N.,
 RA Shimpoo S., Suginoto M., Takeuchi C., Yamada M., Tabata S.;
 RT "Complete genome structure of the thermophilic cyanobacterium
 Thermosynechococcus elongatus BP-1.>";
 RL DNA Res. 9:121-130 (2002).
 DR EMBL; AP005376; BAC09625.1; -.
 DR GO; GO:008658; F:penicillin binding; IEA.
 DR GO; GO:0009273; P:cell wall biosynthesis (sensu Bacteria); IEA.
 DR InterPro; IPR005311; PBP dimer.
 DR InterPro; IPR001460; Transpeptidase.
 DR Pfam; PF00717; PBP dimer; 1.
 DR Pfam; PF00905; Transpeptidase; 1.
 DR GO; GO:0008658; F:penicillin binding; IEA.
 DR GO; GO:0009273; P:cell wall biosynthesis (sensu Bacteria); IEA.
 DR InterPro; IPR005311; PBP dimer.
 DR InterPro; IPR001460; Transpeptidase.
 DR Pfam; PF00717; PBP dimer; 1.
 DR Pfam; PF00905; Transpeptidase; 1.

DR GO; GO:0000155; F:two-component response regulator activity; IEA.
 DR GO; GO:000760; P:sensory perception; IEA.

Query Match 28.8%; Score 60; DB 16; Length 588;
 Best Local Similarity 57.1%; Pred. No. 39; Mismatches 12; Conservative 3; Indels 6; Gaps 0; Gaps 0;

Qy 12 HVVEGLAGELQQLRLEHHIP 32
 DB 442 HVIDGLYDENGQQLRLEHP 462

RESULT 5	
Q6MD1; Q6MD1; PRELIMINARY;	PRT; 139 AA.
AC ID: P1894; H85723.	DR
DT 01-DEC-2001 (TREMBLrel. 19, Created)	DR
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)	DR
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)	DR
DE Hypothetical protein FLW3275.	DR
OS Homo sapiens (Human).	DR
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	DR
OX NCBI - TaxID=9606;	DR
RN [1]	DR
RP SEQUENCE FROM N.A.	DR
RC TISSUE-Spleen;	DR
RA Kawakami B., Sugiyama A., Takemoto M., Sugiyama T., Irie R.,	DR
RA Orsuki T., Sato H., Ota T., Wakamatsu A., Ishii S., Yamamoto J.,	DR
RA Isono Y., Kawai Y., Saito K., Nishikawa T., Kimura K., Yamashita H., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,	DR
RA Watanatsuma M., Murakawa K., Kaneko K., Takahashi-Fujii A., Oshima A.,	DR
RT Suzuki Y., Sugano S., Nagahori K., Masuho Y., Nagai K., Isogai T.; "NEDO human cDNA sequencing Project.";	DR
RL Submitted (Oct-2001) to the EMBL/GenBank/DBJ databases.	DR
KW Hypothetical protein.	DR
SEQUENCE 139 AA; 15014 MW; 3A59D86197B48F19 CRC64;	DR
Query Match 28.4%; Score 59; DB 4; Length 139; Best Local Similarity 44.8%; Pred. No. 11; Matches 13; Conservative 3; Mismatches 11; Indels 2; Gaps 1;	DR
QY 3 EKTIGWRGCHVVEGLAGELBQLRARLEH 31	DR
Db 92 ERATGWRILSHVEQSQGLOATSCR-HH 118	DR
RESULT 6	
Q8XAY8 PRELIMINARY; PRT; 342 AA.	DR
AC ID: P02653; H85723.	DR
DT 01-MAR-2002 (TREMBLrel. 20, Created)	DR
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)	DR
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)	DR
DE Putative transport system permease protein.	DR
GN YDBY OR SF1584 OR S1710!	DR
OS Shigella flexneri	DR
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Shigella.	DR
OC NCBI - TaxID=623;	DR
RN [1]	DR
RP SEQUENCE FROM N.A.	DR
RC STRAIN=301 / Serotype 2a; STRAIN=2457T / ATCC 700930 / Serotype 2a; STRAIN=2457T / ATCC 700930 / Serotype 2a;	DR
RX MEDLINE=22272405; PubMed=12384590; Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H., Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W., Yang J., Zhang J., Wu H., Qu D., Ding J., Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Chen K., Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y., Yu J.;	DR
RA RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity through comparison with genomes of Escherichia coli K12 and O157.";	DR
RA RT Nucleic Acids Res. 30:4422-4441(2002).	DR
RN [2]	DR
RP SEQUENCE FROM N.A.	DR
RC STRAIN=2457T / ATCC 700930 / Serotype 2a; STRAIN=2457T / ATCC 700930 / Serotype 2a; MEDLINE=22590274; PubMed=12704152; Fournier G., Matthew G., Plunkett G., III, Rose D.J., Darling A., Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S., RA RT "Complete genome sequence and comparative genomics of Shigella flexneri serotype 2a strain 2457T.";	DR
RA RT Infect. Immun. 71:2775-2786 (2003). EMBL: AB015180; AAN43171-1; AB015983; AAF17063.1; AB016020; C:membrane; IEA.	DR
RA RT GO:0002215; P:transporter; IEA.	DR
RA DR GO:0006810; P:transport; IBA.	DR
RA DR InterPro; IPR001851; Bac_inmem_trasp.	DR
RC Pfam; PF02653; BPD_transp_2; 1.	DR
RC Complete proteome.	DR
RX SEQUENCE 342 AA; 36396 MW; B5FB60B2A54604B CRC64;	DR
Query Match 28.4%; Score 59; DB 16; Length 342; Best Local Similarity 57.1%; Pred. No. 29; Matches 12; Conservative 2; Mismatches 7; Indels 0; Gaps 0;	DR
QY 8 WRGGHVVVEGLAGELBQLRARL 28	DR
Db 135 WRGKWKVVEGLPALKQSQSAPL 155	DR
RESULT 7	
QY 8 WRGGHVVVEGLAGELBQLRARL 28	DR
Db 135 WRGKWKVVEGLPALKQSQSAPL 155	DR
Query Match 28.4%; Score 59; DB 16; Length 342; Best Local Similarity 57.1%; Pred. No. 29; Matches 12; Conservative 2; Mismatches 7; Indels 0; Gaps 0;	DR
QY 8 WRGGHVVVEGLAGELBQLRARL 28	DR
Db 135 WRGKWKVVEGLPALKQSQSAPL 155	DR
Query Match 28.4%; Score 59; DB 16; Length 342; Best Local Similarity 57.1%; Pred. No. 29; Matches 12; Conservative 2; Mismatches 7; Indels 0; Gaps 0;	DR

Query Match	Score 28.1%; DB 16; Length 274;
Best Local Similarity 38.2%; Pred. No. 26;	Matches 13; Conservative 3; Mismatches 15; Indels 3; Gaps 1;
Matches 13; Conservative 3; Mismatches 15; Indels 3; Gaps 1;	
Qy 2 DEKTTGWRGG--HVEGLAGELQQLRABLEHHP 32	
Db 144 DRRRGWGRKGRETHEAASVUDGTVATLRGDLIHP 177	
RESULT 11	
Q8IU4 PRELIMINARY; PRT; 333 AA.	
ID Q8IU4; AC 01-MAR-2003 (TREMBLrel. 23, Created)	
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)	
DR 01-OCT-2003 (TREMBLrel. 25, Last annotation update)	
DB Hypothetical protein.	
OS Homo sapiens (Human).	
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;	
OC Mammalia; Eutheria; Primates; Catahrini; Hominidae; Homo.	
RA NCBI_TAXID=9606;	
RN [1] SEQUENCE FROM N.A.	
RC TISSUE=Pancreas;	
RA Strausberg R;	
RL Submitted (Nov-2002) to the EMBL/GenBank/DBJ databases.	
DR EMBL; BC040338; AAH4038.1; -.	
DR InterPro; IPR001073; CIG.	
DR InterPro; IPR008160; Collagen.	
DR InterPro; IPR008983; TNF_like.	
DR Pfam; PF00396; CIG; 1.	
DR Pfam; PF01391; Collagen; 3.	
DR PRINTS; PR00007; COMPLNNTCIG.	
DR SMART; SM00110; CIG; 1.	
DR PROSITE; PS01113; CIG; 1.	
DR KW Hypothetical protein.	
SQ SEQUENCE 333 AA; 34649 MW; E9EBF803A034E307 CRC64;	
Query Match 27.9%; Score 58; DB 4; Length 333;	
Best Local Similarity 40.0%; Pred. No. 37;	
Matches 14; Conservative 6; Mismatches 15; Indels 0; Gaps 0;	
Qy 2 DEKTTGWRGGHVVEGLAGELQQLRABLEHHPGOR 36	
Db 90 DQ55RGPCKHKGPKGLAGPMGEKGRLGERGPQGK 124	
RESULT 12	
Q9ZGD9 PRELIMINARY; PRT; 491 AA.	
ID Q9ZGD9; AC 09ZGD9;	
DT 01-MAY-1999 (TREMBLrel. 10, Created)	
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)	
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)	
DB Oxygenase homolog.	
OS Streptomyces cyanogenus.	
OC Streptomyces; Actinobacteridae; Actinomycetales;	
OC Streptomyces; Actinomycetaceae; Streptomyces.	
OC Streptomyces coelicolor.	
OC Streptomyces; Streptomyctaceae; Actinomycetales;	
DR Seeger K.J., Harris D.; Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.	
DR RN [2]	
DR SEQUENCE FROM N.A.	
DR STRAIN=A3 (2);	
DR Cerdeno A.M., Parkhill J., Barrell B.G., Rajandream M.A.; Redenbach M., Kieser H.M., Denapaité D., Eichner A., Cullum J., Kishashi H., Hopwood D.A.; Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.	
DR RN [3]	
DR SEQUENCE FROM N.A.	
DR STRAIN=A3 (2);	
DR MEDLINE-97000351; PubMed=8843436;	
DR Cerdeno A.M., Parkhill J., Barrell B.G., Rajandream M.A.; Redenbach M., Kieser H.M., Denapaité D., Eichner A., Cullum J., Kishashi H., Hopwood D.A.; Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.	
DR RT "A set of ordered cosmids and a detailed genetic and physical map for the 8 Mb Streptomyces coelicolor A3(2) chromosome.";	
DR RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";	
DR Mol. Microbiol. 21:77-96(1996).	
DR RN [4]	
DR SEQUENCE FROM N.A.	
DR STRAIN=A3 (2) / M145; MEDLINE-2199410; PubMed=12000953;	
DR Bentley S.D., Chater K.F., Cerdeno-Taronga A.-M., Challis G.L., Thomson N.R., James D.K.D., Harris D.E., Quail M.A., Kieser H., Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M., Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howard S., Huang C.-H., Kieser T., Larke L., Murphy J., Oliver K., O'Neill S., Rabbitts B., Rajandream M.A., Rutherford K., Rutter S., Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K., Warren T., Wetzorek A., Woodward J., Barrell B.G., Parkhill J., Hopwood D.A.;	
DR SEQUENCE FROM N.A.	
DR STRAIN=S136; MEDLINE=99122695; PubMed=9933932;	
DR SEQUENCE FROM N.A.	
DR [1] Bachthold A.; Cloning and characterization of a gene cluster from Streptomyces cyanogenus S136 probably involved in landomycin biosynthesis.;	
DR FEMS Microbiol. Lett. 170:381-387 (1999).	
DR EMBL; AF080235; ADD13534.1; -.	
DR GO; GO_0004197; F-monooxygenase activity; IFA.	
RESULT 13	
Q9L0L0 PRELIMINARY; PRT; 1161 AA.	
ID Q9L0L0; AC 09L0L0;	
DT 01-OCT-2000 (TREMBLrel. 15, Created)	
DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)	
DR DNA-directed RNA polymerase beta chain.	
DR RPOB OR SCO4554 OR SCB02.26.	
OS Streptomyces coelicolor.	
OC Bacteria; Actinobacteridae; Actinomycetales;	
OC Streptomyces; Streptomyctaceae; Streptomyces.	
OC Streptomyces coelicolor.	
DR NCBI_TAXID=1902;	
DR RN [1]	
DR SEQUENCE FROM N.A.	
DR STRAIN=A3 (2);	
DR RA Seeger K.J., Harris D.; Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.	
DR RN [2]	
DR SEQUENCE FROM N.A.	
DR STRAIN=A3 (2);	
DR MEDLINE-97000351; PubMed=8843436;	
DR RA Cerdeno A.M., Parkhill J., Barrell B.G., Rajandream M.A.; Redenbach M., Kieser H.M., Denapaité D., Eichner A., Cullum J., Kishashi H., Hopwood D.A.; Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.	
DR RN [3]	
DR SEQUENCE FROM N.A.	
DR STRAIN=A3 (2);	
DR MEDLINE-2199410; PubMed=12000953;	
DR Bentley S.D., Chater K.F., Cerdeno-Taronga A.-M., Challis G.L., Thomson N.R., James D.K.D., Harris D.E., Quail M.A., Kieser H., Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M., Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howard S., Huang C.-H., Kieser T., Larke L., Murphy J., Oliver K., O'Neill S., Rabbitts B., Rajandream M.A., Rutherford K., Rutter S., Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K., Warren T., Wetzorek A., Woodward J., Barrell B.G., Parkhill J., Hopwood D.A.;	
DR SEQUENCE FROM N.A.	
DR MEDLINE=99122695; PubMed=9933932;	
DR SEQUENCE FROM N.A.	
DR [1] Bachthold A.; Cloning and characterization of a gene cluster from Streptomyces cyanogenus S136 probably involved in landomycin biosynthesis.;	
DR FEMS Microbiol. Lett. 170:381-387 (1999).	
DR EMBL; AF080235; ADD13534.1; -.	
DR GO; GO_0004197; F-monooxygenase activity; IFA.	
RESULT 14	
Q9L0L0 PRELIMINARY; PRT; 1161 AA.	
ID Q9L0L0; AC 09L0L0;	
DT 01-OCT-2000 (TREMBLrel. 15, Created)	
DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)	
DR DNA-directed RNA polymerase beta chain.	
DR RPOB OR SCO4554 OR SCB02.26.	
OS Streptomyces coelicolor.	
OC Bacteria; Actinobacteridae; Actinomycetales;	
OC Streptomyces; Streptomyctaceae; Streptomyces.	
OC Streptomyces coelicolor.	
DR NCBI_TAXID=1902;	
DR RN [1]	
DR SEQUENCE FROM N.A.	
DR STRAIN=A3 (2);	
DR MEDLINE-97000351; PubMed=8843436;	
DR RA Cerdeno A.M., Parkhill J., Barrell B.G., Rajandream M.A.; Redenbach M., Kieser H.M., Denapaité D., Eichner A., Cullum J., Kishashi H., Hopwood D.A.; Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.	
DR RN [2]	
DR SEQUENCE FROM N.A.	
DR STRAIN=A3 (2);	
DR MEDLINE-2199410; PubMed=12000953;	
DR Bentley S.D., Chater K.F., Cerdeno-Taronga A.-M., Challis G.L., Thomson N.R., James D.K.D., Harris D.E., Quail M.A., Kieser H., Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M., Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howard S., Huang C.-H., Kieser T., Larke L., Murphy J., Oliver K., O'Neill S., Rabbitts B., Rajandream M.A., Rutherford K., Rutter S., Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K., Warren T., Wetzorek A., Woodward J., Barrell B.G., Parkhill J., Hopwood D.A.;	
DR SEQUENCE FROM N.A.	
DR MEDLINE=99122695; PubMed=9933932;	
DR SEQUENCE FROM N.A.	
DR [1] Bachthold A.; Cloning and characterization of a gene cluster from Streptomyces cyanogenus S136 probably involved in landomycin biosynthesis.;	
DR FEMS Microbiol. Lett. 170:381-387 (1999).	
DR EMBL; AF080235; ADD13534.1; -.	
DR GO; GO_0004197; F-monooxygenase activity; IFA.	

!- SUBUNIT: THE ENZYME CONSISTS OF THE SIGMA CHAIN AND THE CORE ENZYME WHICH IS COMPOSED OF 2 ALPHA CHAINS, 1 B β TA CHAIN, AND 1 B β TA' CHAIN (BY SIMILARITY).

-!- SIMILARITY: BELONGS TO THE RNA POLYMERASE BETA CHAIN FAMILY.

CC EMBL: AL939121; CAB77428.1; -.

CC HSSP: Q9KWT7; 1HOM.

DR GO: GO:001677; F:DNA binding; IEA.

DR GO: GO:003899; F:DNA-directed RNA polymerase activity; IEA.

DR GO: GO:0016740; F:transf erase activity; IEA.

DR GO: GO:0005350; P:transcription; IEA.

DR InterPro: IPR007121; RNA_Pol_B.

DR InterPro: IPR007644; RNA_Pol_Rpb2_1.

DR InterPro: IPR007642; RNA_Pol_Rpb2_2.

DR InterPro: IPR007645; RNA_Pol_Rpb2_3.

DR InterPro: IPR007120; RNA_Pol_Rpb2_6.

DR InterPro: IPR007641; RNA_Pol_Rpb2_7.

DR Pfam: PF04561; RNA_Pol_Rpb2_1.

DR Pfam: PF04561; RNA_Pol_Rpb2_2.

DR Pfam: PF04565; RNA_Pol_Rpb2_3.

DR Pfam: PF00562; RNA_Pol_Rpb2_6.

DR Pfam: PF04560; RNA_Pol_Rpb2_7.

DR PROSITE: PS01166; RNA_Pol_BETA; 1.

DR Complete proteome.

KW DNA-directed RNA polymerase; Transcription; Transferase; RNA polymerase; Complete proteome.

SQ 1161 AA; 128494 MW; A9C27B928EE191A8 CRC64;

Query Match 27.9%; Score 58; DB 16; Length 1161; Best Local Similarity 45.5%; Pred. No. 1.5e+02; Matches 15; Conservative 3; Mismatches 13; Indels 2; Gaps 2;

Oy 4 KTGWGRGHVVGVLAGELEQRLARLHRHPOQ 35

Db 227 KALGWTTEQILEEF-GEYESMRATLEKDHQTQGQ 258

RESULT 14

082DQ5 PRELIMINARY; PRT; 1181 AA.

AC 082DQ5; PRELIMINARY; PRT; 1181 AA.

DT 01-JUN-2003 (TREMBLrel. 24, Created)

DT 01-OCT-2003 (TREMBLrel. 24, Last sequence update)

DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)

DE Putative RNA polymerase beta subunit.

GN RPOB OR SAV494.

OS Streptomyces avermitilis.

OC Bacteria; Actinobacteria; Streptomyctaceae; Streptomyces.

OX NCBI_TaxID=33903;

OX NCBI_TaxID=84588;

RESULT 15

07U554 PRELIMINARY; PRT; 229 AA.

AC 07U554; PRELIMINARY; PRT; 229 AA.

DT 01-OCT-2003 (TREMBLrel. 25, Created)

DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)

DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)

DR Hypothetical protein.

GN SYNW855.

OS Synechococcus sp. (strain WH8102).

OC Bacteria; Cyanobacteria; Chroococcales; Synechococcus.

OX NCBI_TaxID=84588;

OX NCBI_TaxID=84588;

RP SEQUENCE FROM N.A.

RX MEDLINE=22825697; PubMed=12917641;

RX Palenik B., Brakhmsia B., Larimer F.W., Land M., Hauser L., Chain P.,

RX Lamerdrin J., Regala W., Allen E.E., McCarren J., Paulsen I.,

RX Dufrene A., Partensky F., Webb E.A., Waterbury J.;

RX "The genome of a mobile marine Synechococcus.;"

RX Nature 424:1037-1042(2003);

RX EMBL: BX59694; CAE0830.1; -.

KW Hypothetical protein; Complete proteome.

SQ 229 AA; 24435 MW; 851B41BB69B319CB CRC64;

Query Match 27.4%; Score 57; DB 16; Length 229; Best Local Similarity 37.5%; Pred. No. 33; Matches 12; Conservative 4; Mismatches 16; Indels 0; Gaps 0;

Oy 5 TTAGWGRGHVVGVLAGELEQRLARLHRHPOQ 36

Db 81 TSGWRETEQAWSAGADWVALIDATORRPEGR 112

Search completed: July 15, 2004, 20:35:48

Job time : 30.509 secs

[1]

SEQUENCE FROM N.A.

RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;

RX MEDLINE=2147403; PubMed=1157248;

RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C., T.,

RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osono T.,

RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;

RT "Genome sequence of an industrial microorganism Streptomyces avermitilis: deducing the ability of producing secondary metabolites";

RT Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).

RN

[2]

SEQUENCE FROM N.A.

RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;

RX MEDLINE=2260306; PubMed=12692562;

RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,

RA Sakaki Y., Hattori M., Omura S.;

RT "Complete genome sequence and comparative analysis of the industrial microorganism Streptomyces avermitilis.;"

RT Nat. Biotechnol. 21:526-531(2003).

DR EMBL: AP005040; BAC7726.1; -.

DR GO; GO:003677; F:DNA binding; IEA.

DR GO; GO:003899; F:DNA-directed RNA polymerase activity; IEA.

DR GO; GO:006350; P:transcription; IEA.

DR InterPro; IPR007121; RNA_Pol_B.

us-10-004-381-25_copy_1_38.rsp

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

On protein - protein search, using SW model

Run on: July 15, 2004, 20:30:05 ; Search time 7.15294 Seconds
(without alignments)
276.623 Million cell updates/sec

Title: US-10-004-381-25_COPY_1_38

Perfect score: 208

Sequence: 1 MDEKTTGMRGGHVVVEGLASLEQLRARLEHHHPQSGREP 38

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Swissprot_42;*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60.5	29.1	177	1 HSILV_RHIL0	Q98ct8 rhizobium 1
2	59	28.4	342	1 YDEY_ECOLI	P05560 corynebacte
3	57	27.4	474	1 TRPC_CORGL	P04448 homo sapien
4	56.5	27.2	404	1 KVB3_HUMAN	Q41229 homo sapien
5	56.5	27.2	1560	1 SMCX_HUMAN	Q8a239 caulobacter
6	55.5	26.7	188	1 HSILV_CAUCR	Q9fr4u streptomyce
7	54	26.0	376	1 ALC_STRCO	Q34949 rattus norv
8	53.5	25.7	404	1 KVB3_RAT	P73950 synochocyst
9	53	25.5	348	1 PLSK_SYN3	P22731 ascaris suu
10	53	25.5	589	1 IFEB_ASCMU	Q8ve31 brucella me
11	52.5	25.2	184	1 HSILV_BRUSU	Q8fy11 brucella su
12	52.5	25.2	184	1 HSILV_BRUSU	P23789 bacillus su
13	52	25.0	113	1 XRE_BACSU	Q20636 caenorhabdi
14	52	25.0	356	1 GB02_CAEEL	Q16352 homo sapien
15	52	25.0	499	1 ALINK_HUMAN	P46660 mus musculus
16	52	25.0	504	1 ALINK_MOUSE	P23565 rattus norv
17	52	25.0	505	1 ALINK_RAT	Q9Y7K2 schizosach
18	52	25.0	2337	1 TOR2_SCHPO	P66886 escherichinia
19	51.5	24.8	181	1 COBU_ECOLI	Q61060 escherichinia
20	51.5	24.8	273	1 GS80_BACSU	P80878 bacillus su
21	51.5	24.8	401	1 KVBL_MOUSE	Q83277 mud musculu
22	51.5	24.8	408	1 KVBL_MOUSE	Q85228 mustella put
23	51.5	24.8	419	1 KVBL_HUMAN	Q44722 homo sapien
24	51.5	24.8	419	1 KVBL_RABBIT	Q9xt31 oryctolagus
25	51.5	24.8	465	1 FXD3_MOUSE	Q61060 mus musculus
26	51	24.5	83	1 RS28_SULSO	Q90095 sulfolobus
27	51	24.5	261	1 YE6V_YEAST	P40078 saccharomyces
28	51	24.5	358	1 ARQB_ARER9	Q9rej9 aeropyrum p
29	51	24.5	429	1 NPX2_MOUSE	O70340 mus musculus
30	51	24.5	431	1 NPX2_HUMAN	P77972 homo sapien
31	51	24.5	432	1 NPX2_RAT	P97738 rattus norv
32	50.5	24.3	149	1 RL13_THEME	Q9x195 thermotoga
33	50.5	24.3	176	1 HSILV_RHIME	Q92tag rhizobium m

RESULT 1		ALIGNMENTS	
HSILV_RHIL0	ID: HSILV_RHIL0	STANDARD:	PRT; 177 AA.
AC: Q98ct8;	DT: 28-FEB-2003	(Rel. 41, Created)	
RT: MEDLINE-2108230;	DT: 28-FEB-2003	(Rel. 41, Last annotation update)	
DE: ATP-dependent protease helv (EC 3.4.25.-).	GN: HSILV OR MJU5007.		
OS: rhizobium loti (Mesorhizobium loti).	OC: Bacteria; Proteobacteria; Alpha-proteobacteria; Rhizobiales;		
OC: Rhizobacteriaceae; Mesorhizobium.	OX: NCBI_TaxID:381;		
RN: [1]	RP: SEQUENCE FROM N.A.		
RC: STRAIN=MARF30309;	RC: MEDLINE=2108230;	PUBLMED=11214968;	
RX: KANEKO_T., NAKAMURA_Y., SATO_S., NAKAZAKI_N., SHIMPO_S., SUGIMOTO_M.,	RX: KISHIDA_Y., KIYOKAWA_C., KOHARA_M., MATSUMOTO_M., MATSUMOTO_A.,	RX: Mochizuki_Y., Nakayama_S., Nakazaki_N., Shimpo_S., Sugimoto_M.,	
RA: KANEKO_T., NAKAMURA_Y., SATO_S., NAKAZAKI_N., SHIMPO_S., SUGIMOTO_M.,	RA: KANEKO_T., NAKAMURA_Y., SATO_S., NAKAZAKI_N., SHIMPO_S., SUGIMOTO_M.,	RA: KANEKO_T., NAKAMURA_Y., SATO_S., NAKAZAKI_N., SHIMPO_S., SUGIMOTO_M.,	
RT: "Complete genome structure of the nitrogen-fixing symbiotic bacterium Mesorhizobium loti." (2000).	RT: "Complete genome structure of the nitrogen-fixing symbiotic bacterium Mesorhizobium loti." (2000).	RT: "Complete genome structure of the nitrogen-fixing symbiotic bacterium Mesorhizobium loti." (2000).	
RL: DNA Res. 7,331-338 (2000).	RL: DNA Res. 7,331-338 (2000).	RL: DNA Res. 7,331-338 (2000).	
CC: -I- FUNCTION: Protease subunit of a proteaseome-like degradation complex (By similarity).	CC: -I- FUNCTION: Protease subunit of a proteaseome-like degradation complex (By similarity).	CC: -I- FUNCTION: Protease subunit of a proteaseome-like degradation complex (By similarity).	
CC: -I- SUBUNIT: A double ring-shaped homohexamer of helv is capped on each side by a ring-shaped helv homohexamer (By similarity).	CC: -I- SUBUNIT: A double ring-shaped homohexamer of helv is capped on each side by a ring-shaped helv homohexamer (By similarity).	CC: -I- SUBUNIT: A double ring-shaped homohexamer of helv is capped on each side by a ring-shaped helv homohexamer (By similarity).	
CC: -I- SUBCELLULAR LOCATION: Cyttoplasmic (By similarity).	CC: -I- SUBCELLULAR LOCATION: Cyttoplasmic (By similarity).	CC: -I- SUBCELLULAR LOCATION: Cyttoplasmic (By similarity).	
CC: -I- SIMILARITY: Belongs to peptidase family T1B. Helv subfamily.	CC: -I- SIMILARITY: Belongs to peptidase family T1B. Helv subfamily.	CC: -I- SIMILARITY: Belongs to peptidase family T1B. Helv subfamily.	
CC: This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or_send_an_email_to_licensee@isb-sib.ch).	CC: This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or_send_an_email_to_licensee@isb-sib.ch).	CC: This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or_send_an_email_to_licensee@isb-sib.ch).	
CC: DR: EMBL; AP003005; BAB51533.1; -.	CC: DR: MEROPS; T01-006; -.	CC: DR: MEROPS; T01-006; -.	
CC: DR: MAPDB; MP_00248; -; 1.	CC: DR: MAPDB; MP_00248; -; 1.	CC: DR: MAPDB; MP_00248; -; 1.	
CC: DR: InterPro; IPR001353; Peptidase_T1L.	CC: DR: InterPro; IPR001353; Peptidase_T1L.	CC: DR: InterPro; IPR001353; Peptidase_T1L.	
CC: DR: Pfam; PF00227; proteaseome_1.	CC: DR: Pfam; PF00227; proteaseome_1.	CC: DR: Pfam; PF00227; proteaseome_1.	
CC: KW: Hydrolyse; Protease; Threonine protease; Complete protease.	CC: KW: Hydrolyse; Protease; Threonine protease; Complete protease.	CC: KW: Hydrolyse; Protease; Threonine protease; Complete protease.	
CC: FT: ACT_SITE; 4 BY SIMILARITY	CC: FT: ACT_SITE; 4 BY SIMILARITY	CC: FT: ACT_SITE; 4 BY SIMILARITY	
CC: SEQUENCE: 177 AA; 18986 MW; 62F3D3DC6ACB5F05 CRC64;	CC: SEQUENCE: 177 AA; 18986 MW; 62F3D3DC6ACB5F05 CRC64;	CC: SEQUENCE: 177 AA; 18986 MW; 62F3D3DC6ACB5F05 CRC64;	
CC: Q: Query Match 29.1%; score 60.5; DB 1; length 177; Best Local Similarity 41.9%; pred. No. 0.9; Matches 13; Conservative 6; Mismatches 5; Indels 7; Gaps 1;	CC: Q: Query Match 29.1%; score 60.5; DB 1; length 177; Best Local Similarity 41.9%; pred. No. 0.9; Matches 13; Conservative 6; Mismatches 5; Indels 7; Gaps 1;	CC: Q: Query Match 29.1%; score 60.5; DB 1; length 177; Best Local Similarity 41.9%; pred. No. 0.9; Matches 13; Conservative 6; Mismatches 5; Indels 7; Gaps 1;	
CC: Q: 9 RGSHGVVGLAGE-----LEQIARLEHHHP 32	CC: Q: 9 RGSHGVVGLAGE-----LEQIARLEHHHP 32	CC: Q: 9 RGSHGVVGLAGE-----LEQIARLEHHHP 32	
CC: Q: 9 Q92tag rhizobium m	CC: Q: 9 Q92tag rhizobium m	CC: Q: 9 Q92tag rhizobium m	

entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC

CC

DR

DR

KW

DR

entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

Query Match 26.0%; Score 54; DB 1; Length 376; Best Local Similarity 44.4%; Pred. No. 14; Matches 12; Conservative 1; Mismatches 12; Indels 2; Gaps 1;

Qy 10 GGSIVVEGLAGELQ-LRARLEHHPQG 34

Db 171 GGHANGFAVSAEQRFTHRLKOHPDG 197

RESULT 8

KV3 RAT STANDARD; PRT; 404 AA.

ID KV3 RAT

AC 063454; 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)

DE Voltage-gated potassium channel beta-3 subunit (K⁺ channel beta-3 subunit)

DE (KV-beta-3) (RCK beta3).

GN KCNA3 OR CKBETAB3.

OS Rattus norvegicus (Rat).

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OC NCBI_TAXID=10116;

OX [1]

RN SEQUENCE FROM N.A., FUNCTION, AND TISSUE SPECIFICITY.

RC TISSUE=Brain cortex; MEDLINE=96140552; Pubmed=8549760;

RX Heinemann S.H., Retig J., Wunder F., Pongs O.; "Molecular and functional characterization of a rat brain Kv beta 3 potassium channel subunit"; FEBS Lett. 377:383-386(1995).

CC -!- FUNCTION: Accessory potassium channel protein which modulates the activity of the pore-forming alpha subunit. Alters the functional properties of Kv1.4 but not Kv1.1 or Kv1.5.

CC -!- SUBUNIT: Forms heteromultimeric complex with alpha subunits.

CC -!- SUBCELLULAR LOCATION: Cyttoplasmic (Potential)

CC -!- TISSUE SPECIFICITY: Predominantly expressed in brain. Strongest expression in olfactory bulb and thalamic nuclei. Not detected in heart, spleen, lung, liver, skeletal muscle, kidney and testis.

CC -!- DOMAIN: Alteration of functional properties of alpha subunit is mediated through N-terminal domain of beta subunit (probable).

CC -!- SIMILARITY: Belongs to the shaker potassium channel family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC DR EMBL; D90911; BRA18017.1; -.

CC DR PIR; S75456; S75456.

CC DR HAMAP; MF_00019.1; -.

CC DR InterPro; IPR003664; FAI synthesis.

CC DR Pfam; PF02504; FAI synthesis; 1.

CC DR ProDom; PD006974; FAI synthesis; 1.

CC DR TIGRFAMs; TIGR00182; PLS; 1.

CC KW Fatty acid biosynthesis; Phospholipid biosynthesis; Complete proteome.

CC SEQUENCE 348 AA; 37285 MW; 12C07CCF971A72 CRC64;

Qy 2 DEKTTGPRG-GHVVVVEGLAGELQ-LRARLEHHPQG 34

Db 255 EBLPRGWRKGKUAI1--LAPNLKRIKRQVDHAEHG 287

RESULT 10

IPFB ASCSU STANDARD; PRT; 589 AA.

ID IPFB ASCSU

AC P23731; 01-NOV-1991 (Rel. 20, Created)

DT 01-NOV-1991 (Rel. 20, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Intermediate filament protein B (IF-B).
 OS Ascaris suum (PIG roundworm) (Ascaris lumbricoides).
 OC
 OC Ascarididae; Ascaris.
 RN [1] TAXID=625;
 RT proteins; sequence characterization of two muscle proteins of a
 RT nematode.";
 RX MEDLINE=90059912; PubMed=2583097;
 RA Weber K., Plessmann U., Ulrich W.;
 RT "Cytoplasmic intermediate filament proteins of invertebrates are
 closer to nuclear lamins than are vertebrate intermediate filament
 CC SEQUENCE.
 CC -!- SUBUNIT: A and B can form homopolymers.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -!- TISSUE SPECIFICITY: Giant body muscle cells.
 CC -!- SIMILARITY: Belongs to the intermediate filament family.
 DR InterPro: IPR013221; IF tail_C.
 DR Pfam: PF00338; filament_1.
 DR Pfam: PF00332; IF tail_1.
 DR PROSITE: PS00226; IF_1.
 KW Intermediate filament; Coiled coil.
 FT DOMAIN 1 84 HEAD.
 FT DOMAIN 85 433 ROD.
 FT DOMAIN 434 589 TAIL.
 FT DOMAIN 85 116 COIL_1A.
 FT DOMAIN 117 130 LINKER_1.
 FT DOMAIN 131 268 COIL_1B.
 FT DOMAIN 269 285 COIL_2.
 SQ SEQUENCE 589 AA; 67694 MW; 5EF79F6FOAFD6 CRC64;

Query Match 25.5%; Score 53; DB 1; Length 589;
 Best Local Similarity 48.0%; Pred. No. 31; Mismatches 9; Indels 2; Gaps 1;
 Matches 12; Conservative 2; Mismatches 5; Indels 7; Gaps 1;

Qy 3 EKUTGNGGGHVVYEGLAGLEQLRAR 27
 Db 182 EKINQWQ-HATEQDAQSBLEMLRAR 204

RESULT 11
 HSLV_BRUME STANDARD; PRT; 184 AA.
 ID HSLV_BRUME STANDARD; PRT; 184 AA.
 AC Q8YE31;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 10-OCT-2003 (Rel. 41, Last sequence update)
 DE ATP-dependent protease hslv (EC 3.4.25.-).
 OS HSLV OR BR2080.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Brucellaceae; Brucella.
 OC NCBI_TaxID=29461;
 RN [1] TAXID=29461;
 RP SEQUENCE FROM N.A.
 RC STRAIN=1330 / Bovar 1;
 RX MEDLINE=2247741; PubMed=12271122;
 RA Paulsen I.T., Seshadri R., Nelson K.E., Eisen J.A., Heidelberg J.F.,
 RA Read T.D., Dodson R.J., Umayam L., Brinkac L.M., Beanan M.J.,
 RA Daugherty S.C., Deboy R.T., Durkin A.S., Kolonay J.F., Madupu R.,
 RA Nelson W.C., Ayodeji B., Kraul M., Shetty J., Malek J., Van Aken S.E.,
 RA Riedmiller S., Tettelin H., Gill S.R., White O., Salzberg S.L.,
 RA Hoover D.L., Lindler L.E., Halling S.M., Boyle S.M., Fraser C.M.,
 RA "The Brucella suis Genome reveals fundamental similarities between
 RT animal and plant pathogens and symbionts.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:13148-13153 (2002).
 CC -!- FUNCTION: Protease subunit of a proteasome-like degradation
 CC complex (By similarity).
 CC -!- SUBUNIT: A double ring-shaped homohexamer of hslv is capped on
 CC each side by a ring-shaped hslv homohexamer (By similarity).
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -!- SIMILARITY: Belongs to peptidase family T1B. Hslv subfamily.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; AE014495; AAC30970.1; -.
 DR HAMAP; MF_00248; -; 1.

DR InterPro; IPR001353; Peptidase_T1.
 DR Pram; PRO0227; protease; 1.
 KW Hydrolase; Protease; Threonine protease; Complete proteome.
 FT ACT SITE 12 12 BY SIMILARITY.
 SEQUENCE 184 AA; 19809 MW; 1B5B1C68F4A2FEE2 CRC64;
 Query Match 25.2%; Score 52.5; DB 1; Length 184;
 Best Local Similarity 41.4%; Pred. No. 10;
 Matches 12; Conservative 5; Mismatches 5; Indels 7; Gaps 1;
 OY 11 GHVVEGLACE-----LQQQLARLHHP 32
 DB 51 GNVVAGFAGATAFTLLEQLKLEQYP 79

RESULT 13
 XRE_BASU XRE_BASU STANDARD; PRT; 113 AA.
 AC P23789;
 DT 01-NOV-1991 (Rel. 20, Created)
 DT 01-NOV-1991 (Rel. 20, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE HTH-type transcriptional regulator xre (putative pbX repressor).
 GN XRE OR BSU2510.
 OS *Bacillus subtilis*.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; *Bacillus*.
 OX NCBI_TAXID=1423;
 RN [1] MEDLINE=9102505; PubMed=2125016;
 RX Wood H.E., Devine K.M., McConnell D.J.;
 RT 'Characterisation of a repressor gene (*xre*) and a a.
 RT temperature-sensitive allele from the *Bacillus subtilis* prophage,
 RT PBXN.';
 RL Gene 96:83-88(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168 / S0113;
 RX MEDLINE=9434963; PubMed=8083174;
 RA McDonnell G.E., Wood H., Devine K.M., McConnell D.J.;
 RT 'Genetic control of bacterial suicide: regulation of the induction of
 PBXN in *Bacillus subtilis*'.
 J. Bacteriol. 176:5820-5830(1994).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RA Krogh S., O'Reilly M., Nolan N., Devine K.M.;
 RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=98144033; PubMed=9384377;
 RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchart S.,
 RA Boulle R., Bourquier L., Brans A., Braun M., Brignall S.C., Bron S.,
 RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
 RA Choi S.K., Codani J.J., Connerter I.F., Cummings N.J., Daniel R.A.,
 RA Denizot F., Devine K.M., Duschettoff A., Ehrlich S.D., Eumerson P.T.,
 RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
 RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
 RA Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
 RA Guiseppi G., Guy B.J., Haga K., Hailech G., Harwood C.R., Heaut A.,
 RA Hilbert H., Hollappel S., Hosono S., Hulio M.F., Itaya M., Jones L.,
 RA Joris B., Karamata D., Kasahara Y., Klaer-Borchard M., Klein C.,
 RA Kobayashi Y., Koetter P., Konigstein G., Krogh S., Kumano M.,
 RA Kurita K., Lapius A., Lardinois S., Laufer J., Lazarevic V.,
 RA Lee S.M., Levine A., Liu H., Masuda S., Maeli C., Medigue C.,
 RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
 RA Noone D., O'Reilly M., Ogawa K., Ogiwa A., Oudega B., Park S.H.,
 RA Parro V., Pohl T.M., Porteboeuf D., Porwollik S., Prescott A.M.,
 RA presecan E., Pujo P., Purnelle B., Raport G., Ray M., Reynolds S.,
 RA Rieger M., Rivoira C., Rocha E., Roche B., Rose M., Sadie Y.,
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scuffone F.,

RA Sekiguchi J., Sekowska A., Seror S.J., Serrur P., Shin B.S., Soldo B.,
 RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
 RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
 RA Nobato V., Uchiyama S., Vandembol M., Vannier F., Vassarotti A.,
 RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzsaegeger T.,
 RA Winters P., Wipat A., Yamamoto H., Yamane K., Yamamoto K., Yata K.,
 RA Yoshida K., Yosikawa H., Zimstein E., Yosikawa H., Danchin A.,
 RT "The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*".
 RL Nature 390:249-256 (1997).
 CC -!- FUNCTION: PROBABLE REPRESSOR. NECESSARY FOR THE MAINTENANCE OF THE LYSGENIC STATE.
 CC -!- MISCCELLANEOUS: THE DEFECTIVE PROPHAGE OF *BACILLUS SUBTILIS* 168,
 CC PBXN, IS A CHROMOSOMALLY BASED ELEMENT WHICH ENCODES A NON-INFECTIONOUS PHAGE-LIKE PARTICLE WITH BACTERICIDAL ACTIVITY.
 CC PBXN IS INDUCED BY AGENTS WHICH ELICIT THE SOS RESPONSE.
 CC -!- SIMILARITY: Contains 1 HTH cro/CI-type DNA-binding domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation on the European Bioinformatics Institute. There are no restrictions on its use. By non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC DR EMBL; M36478; AAA22941;
 CC DR EMBL; M36477; AAA22893.1; --;
 CC DR EMBL; 234287; CAA84042.1; --;
 CC DR EMBL; 270177; CAA94052.1; --;
 CC DR EMBL; Z9910; CAB13108.1; --;
 CC DR PIR; JQ0810; JQ0810.
 CC DR Subtilist; BG10994; xre.
 CC DR InterPro; IPR01387; HTH_3.
 CC DR Pfam; PF01381; HTH_3.1.
 CC DR SMART; SM00530; HTH_XRE; 1.
 CC DR PROSITE; P850943; HTH_CROCI; 1.
 CC DR TRANSCRIPTION REGULATION; REPRESSOR; DNA-BINDING; COMPLETE PROTEOME.
 CC DOMAIN 6 58 H-T-H MOTIF (BY SIMILARITY).
 CC DNA BIND 15 34 H-T-H MOTIF (BY SIMILARITY).
 CC FT VARIANT 4 4 G -> S (IN TEMPERATURE-SENSITIVE ALLELE).
 CC FT VARIANT 19 19 A -> V (IN TEMPERATURE-SENSITIVE ALLELE).
 CC FT VARIANT 78 78 L -> V (IN TEMPERATURE-SENSITIVE ALLELE).
 CC SQ SEQUENCE 113 AA; 13258 MW; 23AF5C9DB3D2B31A CRC64;
 CC
 Query Match 25.0%; Score 52; DB 1; Length 113;
 Best Local Similarity 36.7%; Pred. No. 7.2;
 Matches 11; Conservative 4; Mismatches 15; Indels 0; Gaps 0;
 OY 9 RGGHVVEGLACE-----LQQQLARLHHPQQCQEP 38
 DB 10 RGRKTOEELASHIGVSRARYSHYENGSE 39

RESULT 14
 GBB2_CAEEL GBB2_CAEEL STANDARD; PRT; 356 AA.
 AC Q20635;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Guanine nucleotide-binding protein beta subunit 2.
 RN GBB2 OR F52A.2.
 OS Caenorhabditis elegans.
 OC Bukaryota; Metazoa; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditida; Palaearctida; Caenorhabditis.
 OC NCBI_TAXID=6239;
 RN [1] SEQUENCE FROM N.A., AND INTERACTIONS.
 RC STRAIN=Bristol N2;
 RX MEDLINE=21231612; PubMed=1133232;
 RA "The G protein beta subunit gpb-2 in *Caenorhabditis elegans* regulates

RT the *G*(*o*)alpha-*G*(*q*)alpha signaling network through interactions with
 RT the regulator of *G* protein signaling proteins *egl-10* and *eat-16*.";
 RL Genetics 158:221-235(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Bristol N2;
 RA Wilkinson J.;
 RL Submitted (APR-1996) to the EMBL/GenBank/DDBJ databases.
 CC -!- FUNCTION: Ghanine nucleotide-binding proteins (G proteins) are
 involved as a modulator or transducer in various transmembrane
 signaling systems. The beta and gamma chains are required for the
 G-protein activity, for replacement of GDP by GTP, and for G protein-
 effector interaction.
 CC -!- SUBUNIT: G proteins are composed of 3 units, alpha, beta and
 gamma. Interacts with *gao-1*, *eat-16*, *egl-10*, *egl-30*.
 CC -!- SIMILARITY: Contains 7 WD repeats.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF291847; AAK5964.1; -.
 DR EMBL; Z71263; CA95824.1; -.
 DR PIR; T22478; T22478.
 DR HSSP; P04901; I1BG.
 DR WormRep; F52A8.2; CE17845.
 DR InterPro; IPR01632; Gprotein_B.
 DR InterPro; IPR001680; WD40.
 DR Pfam; PF00400; WD40; WD40.
 DR PRINTS; PR00319; GPROTEINB.
 DR PRODOM; PR00320; GPROTEINRPT.
 DR PRODOM; PD00018; WD0; 3.
 DR SMART; SMD0320; WD40; 7.
 DR PROSITE; PS00678; WD_REPEATS_1; 3.
 DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
 KW Transducer; Repeat; WD repeat; Multigene family.
 FT REPEAT 66 WD 96 WD 1.
 FT REPEAT 108 138 WD 2.
 FT REPEAT 154 184 WD 3.
 FT REPEAT 196 228 WD 4.
 FT REPEAT 240 270 WD 5.
 FT REPEAT 284 314 WD 6.
 FT REPEAT 326 356 WD 7.
 SQ SEQUENCE 356 AA; 39453 MW; 293AA6DC317D832 CRC64;
 Query Match 25.0%; Score 52; DB 1; Length 356;
 Best Local Similarity 40.7%; P-Value: 2.4;
 Matches 11; Conservative 6; Mismatches 10; Indels 0; Gaps 0;
 QY 3 EKTTGWMGGHVVLEGAGLEDLQARLVE 29
 QY :|||:|||:|||:|||:|||:
 Db 6 QPPTTEKSEYLEOLANAEELRKKLD 32
 RESULT 15
 AINX_HUMAN STANDARD; PRT; 499 AA.
 ID -AINX_HUMAN STANDARD; PRT; 499 AA.
 AC Q16357; Q8RC5;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Alpha-internexin (Alpha-Inx) (66 kDa neurofilament protein)
 DE (Neurofilament-66) (NP-66).
 INA.
 OS Homo sapiens (Human);
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarhini; Hominoidea; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=fetal brain;
 RX MEDLINE=95287809; PubMed=7769995;
 RA Chan S.-O., Chiu F.-C.;
 RC "Cloning and developmental expression of human 66 kd neurofilament
 protein.,"
 RA Brain Res., Mol. Brain Res. 29:177-184 (1995).
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Sheinmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Butow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marsina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshimuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquaiello N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.C., Muzny D.M., Soderren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green S.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska M., Smailus D.E.,
 RA Schnarch A., Schein J.B., Jones S.J.M., Maira M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences.,"
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 CC
 CC -!- FUNCTION: CLASS-IV MUSCULAR INTERMEDIATE FILAMENT THAT IS ABLE TO
 CC SELF-ASSEMBLE. IT IS INVOLVED IN THE MORPHOGENESIS OF NEURONS. IT
 CC MAY FORM AN INDEPENDENT STRUCTURAL NETWORK WITHOUT THE INVOLVEMENT
 CC OF OTHER NEUROFILAMENTS OR IT MAY COOPERATE WITH NF-L TO FORM THE
 CC FILAMENTOUS BACKBONE TO WHICH NF-M AND NF-H ATTACH TO FORM THE
 CC CROSS-BRIDGES.
 CC
 CC -!- TISSUE SPECIFICITY: FOUND PREDOMINANTLY IN ADULT CNS.
 CC -!- DEVELOPMENTAL STAGE: EXPRESSED IN BRAIN AS EARLY AS THE 16TH OF
 CC GESTATION, AND INCREASED RAPIDLY AND REACHED A STEADY STATE LEVEL
 CC BY THE 18TH WEEK OF GESTATION.
 CC
 CC -!- SIMILARITY: Belongs to the intermediate filament family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; BC006359; AAB34482.1; -.
 DR EMBL; BC006359; AAB06359.1; -.
 DR PIR; I52558; I52658.
 DR Genew; HGNC; 6057; INA.
 DR GO; GO:0005883; C:neurofilament; TAS.
 DR GO; GO:0005200; F:structural constituent of cytoskeleton; TAS.
 DR InterPro; IPR06821; Filament_head.
 DR InterPro; IPR01664; IF.
 DR Pfam; PF00038; Filament; 1.
 DR Pfam; PF04732; filament_head; 1.
 DR PROSITE; PS00226; IF; 1.
 KW Intermediate filament; Coiled coil; Neurone.
 FT DOMAIN 1 87 HEAD.
 FT DOMAIN 88 408 ROD.
 FT DOMAIN 409 499 TAIL.
 FT DOMAIN 88 129 COIL 1A.
 FT DOMAIN 130 142 LINKER 1.
 FT DOMAIN 143 238 COIL 1B.
 FT DOMAIN 239 262 LINKER 2.
 FT DOMAIN 263 408 COIL 2.

FT DOMAIN 449 454 POLY-GLU.
 FT CONFLICT 37 41 GFPSQ -> ASVE (IN REF. 1).
 FT CONFLICT 67 67 R -> A (IN REF. 1).
 FT CONFLICT 92 92 T -> S (IN REF. 1).
 FT CONFLICT 128 132 ALQR -> RCDT (IN REF. 1).
 FT CONFLICT 141 141 E -> Q (IN REF. 1).
 FT CONFLICT 147 152 LDRRA -> PRHP (IN REF. 1).
 FT CONFLICT 191 198 GARRAKA -> RRARLCK (IN REF. 1).
 FT CONFLICT 244 244 A -> R (IN REF. 1).
 FT CONFLICT 263 263 S -> A (IN REF. 1).
 FT CONFLICT 301 301 S -> T (IN REF. 1).
 FT CONFLICT 310 318 ER -> DQ (IN REF. 1).
 FT CONFLICT 318 318 MISSING (IN REF. 1).
 SQ SEQUENCE 499 AA; 55390 MW; 4C972/64E9B68D3E CRC64;

Query Match 25.0%; Score 52; DB 1; Length 499;
 Best Local Similarity 47.6%; Prod. No. 35;
 Matches 10; Conservative 6; Mismatches 5; Indels 0;
 Gaps 0;

QY 15 EGLAGELEQURABLEHHFQ3Q 35
 Db 169 DGLAEEVQRARCEEEGR 189

Search completed: July 15, 2004, 20:34:29
 Job time : 8.15294 SECs

This Page Blank (uspto)

Copyright (c) 1993 - 2004 Compugen Ltd.	GenCore version 5.1.6				
Run on: July 15, 2004, 20:31:20 ; Search time 10.2824 Seconds	(without alignments)				
Scoring table: BLOSUM62	355.490 Million cell updates/sec				
Title: US-10-004-381-25_COPY_1_38					
perfect score: 208					
Sequence: 1 MDEKTTGWRGGHVVVEGLAGBLEQLRARLBHHPOQOREP 38					
Scored:					
Total number of hits satisfying chosen parameters: 283366					
Minimum DB seq length: 0					
Maximum DB seq length: 200000000					
Post-processing: Minimum Match 0% Listing first 45 summaries	Maximum Match 100%				
Database : PIR 78:*					
1: pir1:*					
2: pir2:*					
3: pir3:*					
4: pir4:*					
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.					
SUMMARIES					
Result No.	Score	Query Match	Length	DB ID	Description
1	59	28.4	342	2	BB4905 probable sugar transport permease protein b1514 - Escherichia coli (strain K-12)
2	59	28.4	342	2	BB4905 probable sugar transport permease protein b1514 - Escherichia coli (strain K-12)
3	59	28.4	342	2	BB4905 probable sugar transport permease protein b1514 - Escherichia coli (strain K-12)
4	57	27.4	474	1	BB4905 probable sugar transport permease protein b1514 - Escherichia coli (strain K-12)
5	57	27.4	663	2	BB4905 probable sugar transport permease protein b1514 - Escherichia coli (strain K-12)
6	56.5	27.2	1560	2	BB4905 probable sugar transport permease protein b1514 - Escherichia coli (strain K-12)
7	55.5	26.7	1543	1	BB4905 probable sugar transport permease protein b1514 - Escherichia coli (strain K-12)
8	54	26.0	321	2	BB4905 probable sugar transport permease protein b1514 - Escherichia coli (strain K-12)
9	54	26.0	1156	2	BB4905 probable sugar transport permease protein b1514 - Escherichia coli (strain K-12)
10	53.5	25.7	404	2	BB4905 probable sugar transport permease protein b1514 - Escherichia coli (strain K-12)
11	53.5	25.7	554	2	BB4905 probable sugar transport permease protein b1514 - Escherichia coli (strain K-12)
12	53	25.5	261	2	BB4905 probable sugar transport permease protein b1514 - Escherichia coli (strain K-12)
13	53	25.5	348	1	BB4905 probable sugar transport permease protein b1514 - Escherichia coli (strain K-12)
14	53	25.5	495	2	BB4905 probable sugar transport permease protein b1514 - Escherichia coli (strain K-12)
15	53	25.5	589	2	BB4905 probable sugar transport permease protein b1514 - Escherichia coli (strain K-12)
16	53	25.5	947	2	BB4905 probable sugar transport permease protein b1514 - Escherichia coli (strain K-12)
17	52.5	25.2	184	2	BB4905 probable sugar transport permease protein b1514 - Escherichia coli (strain K-12)
18	52.5	25.2	392	2	BB4905 probable sugar transport permease protein b1514 - Escherichia coli (strain K-12)
19	52.5	25.2	744	2	BB4905 probable sugar transport permease protein b1514 - Escherichia coli (strain K-12)
20	52	25.0	113	2	BB4905 probable sugar transport permease protein b1514 - Escherichia coli (strain K-12)
21	52	25.0	356	2	BB4905 probable sugar transport permease protein b1514 - Escherichia coli (strain K-12)
22	52	25.0	376	2	BB4905 probable sugar transport permease protein b1514 - Escherichia coli (strain K-12)
23	52	25.0	494	2	BB4905 probable sugar transport permease protein b1514 - Escherichia coli (strain K-12)
24	52	25.0	504	2	BB4905 probable sugar transport permease protein b1514 - Escherichia coli (strain K-12)
25	52	25.0	505	2	BB4905 probable sugar transport permease protein b1514 - Escherichia coli (strain K-12)
26	52	25.0	636	2	BB4905 probable sugar transport permease protein b1514 - Escherichia coli (strain K-12)
27	52	25.0	2337	2	BB4905 probable sugar transport permease protein b1514 - Escherichia coli (strain K-12)
28	51.5	24.8	571	2	BB4905 probable sugar transport permease protein b1514 - Escherichia coli (strain K-12)
29	51.5	24.8	181	2	BB4905 probable sugar transport permease protein b1514 - Escherichia coli (strain K-12)
ALIGNMENTS					
RESULT 1					
BB4905	probable sugar transport permease protein b1514 - Escherichia coli (strain K-12)				
C;Species: Escherichia coli					
C;Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 01-Mar-2002					
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co					
A.; Rose, D.J.; Mau, B.; Shao, Y.					
Science 277, 1453-1462, 1997					
A;Title: The complete genome sequence of Escherichia coli K-12.					
A;Reference number: BB4905; MVID:97426617; PMID:9278503					
A;Accession: BB4905					
A;Status: nucleic acid sequence not shown; translation not shown					
A;Molecule type: DNA					
A;Residues: 1-442 					
A;Cross-references: GB:AB00249; GB:U00096; MVID:91787790; PIDN: AAC74597.1; PID:91787793,					
C;Experimental source: strain K-12, substrain MG1655					
C;Superfamily: 1-arabinose transport system permease arah					
C;Keywords: transmembrane protein					
F:11-27/Domain: transmembrane #status predicted <TM1>					
F:44-60/Domain: transmembrane #status predicted <TM2>					
F:65-81/Domain: transmembrane #status predicted <TM3>					
F:88-104/Domain: transmembrane #status predicted <TM4>					
F:107-123/Domain: transmembrane #status predicted <TM5>					
F:155-171/Domain: transmembrane #status predicted <TM6>					
F:208-224/Domain: transmembrane #status predicted <TM7>					
F:247-263/Domain: transmembrane #status predicted <TM8>					
F:266-282/Domain: transmembrane #status predicted <TM9>					
F:291-307/Domain: transmembrane					
Query Match Best Local Similarity 57.1%; Pred. No. 5.4%; Matches 12; Conservative 2; Mismatches 7; Indels 0; Gaps 0; Gaps 0;					
DB 135 WIGKWMIEGLPAELKQISPL 155					
RESULT 2					
A90894					
probable transport system permease protein bcs2121 [imported] - Escherichia coli (strain					
C;Species: Escherichia coli					
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001					
C;Accession: A90894					
R;Hayashi, T.; Makino, K.; Ohishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.					
Gasawara, N.; Yasunaga, T.; Kumar, S.; Shiba, T.; Hattori, M.; Shinagawa, H.					
DNA Res. 8, 11-22, 2001					
A;Title: Complete genome sequence of enteropathogenic Escherichia coli O157:H7 and gen					
A;Reference number: A99629; MVID:21156231; PMID:11258796					
A;Accession: A90894					
A;Status: preliminary					

A; Molecule type: DNA
 A; Residues: 1-342 <HAY>
 A; Cross-references: GB:BA000007; PIDN:BAR35544.1; PID:913361587; GSPDB:GN00154
 A; Experimental source: strain O157:H7, Substrain RIMD 0509952
 C; Genetics:
 A; Gene: EC2121
 C; Superfamily: 1-arabinose transport system permease arah
 Query Match 28.4%; Score 59; DB 2; Length 342;
 Best Local Similarity 57.1%; Pred. No. 5, 4; Mismatches 7; Indels 0; Gaps 0;
 Matches 12; Conservative 2; Mismatches 7; Indels 0; Gaps 0;
 Oy 8 WRGCHHVVEGLAGLEQLRRL 28
 Db 135 WNGGKWKIEGLPAELKQLSAPL 155

RESULT 3

H85723 probable transport system permease protein ydeY [imported] - Escherichia coli (strain O1) probable transport system permease protein ydeY [imported] - Escherichia coli (strain O1)
 C; Species: Escherichia coli
 C; Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
 C; Accession: H85723
 R; Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew, I.; Grotnbeck, E.J.; Davis, N.W.; Lim, A.; Bimantara, B.; Potamitis, K.; Apodaca, Nature 409, 529-533, 2001
 A; Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
 A; Reference number: A85480; MUID:21074935; PMID:11206551
 A; Accession: H85723
 A; Status: preliminary
 A; Molecule type: DNA
 A; Residues: 1-342 <STO>
 A; Experimental source: strain O157:H7, substrain EDL933
 C; Genetics:
 A; Gene: ydeY
 C; Superfamily: 1-arabinose transport system permease arah
 Query Match 28.4%; Score 59; DB 2; Length 342;
 Best Local Similarity 57.1%; Pred. No. 5, 4; Mismatches 7; Indels 0; Gaps 0;
 Matches 12; Conservative 2; Mismatches 7; Indels 0; Gaps 0;
 Oy 8 WNGGKWKIEGLPAELKQLSAPL 28
 Db 135 WNGGKWKIEGLPAELKQLSAPL 155

RESULT 4

B24723 trpC Protein - Corynebacterium glutamicum
 C; Species: Corynebacterium glutamicum
 C; Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 C; Accession: B24723
 R; Matcui, K.; Sano, K.; Ohtsubo, E.
 Nucleic Acids Res. 14, 10113-10114, 1986
 A; Title: Complete nucleotide and deduced amino acid sequences of the *Brevibacterium lact*
 A; Reference number: A93606; MUID:87117512; PMID:3808947
 A; Accessions: B24723
 A; Residues: 1-474 <NAT>
 A; Molecule type: DNA
 A; Cross-references: GB:X04960; NID:939591; PIDN:CAA28626.1; PID:939596
 C; Genetics:
 A; Gene: trpC
 C; Superfamily: trpC-trpP bifunctional enzyme; trpC homology; trpP homology
 F; 9-28/Domain: trpC homology <TRC>
 F; 264-460/Domain: trpP homology <TRP>
 Query Match 27.4%; Score 57; DB 1; Length 474;
 Best Local Similarity 55.6%; Pred. No. 14; Mismatches 10; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
 Oy 13 WVEGLAGLEQLRRL 30

RESULT 5

DB3545 probable helicase PA0799 [imported] - *Pseudomonas aeruginosa* (strain PA01)
 C; Species: *Pseudomonas aeruginosa*
 C; Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
 C; Accession: DB3545
 R; Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrener, P.; Hickey, M.J.; Br adman, S.; Yuan, Y.; Brady, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, J.; Lory, S.; Olson, M.V.
 Nature 406, 959-964, 2000
 A; Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pathogen
 A; Reference number: A82950; MUID:2043733; PMID:10984043
 A; Accession: D83545
 A; Status: preliminary
 A; Molecule type: DNA
 A; Residues: 1-663 <STO>
 A; Cross-references: GB:AE004515; GB:AE004091; NID:9946687; PIDN:AAG04188.1; GSPDB:GN00111
 A; Experimental source: Strain PA01
 C; Genetics:
 A; Gene: PA0799
 Query Match 27.4%; Score 57; DB 2; Length 663;
 Best Local Similarity 36.1%; Pred. No. 20; Mismatches 13; Conservative 3; MiMismatches 20; Indels 0; Gaps 0;
 Matches 13; Conservative 3; MiMismatches 20; Indels 0; Gaps 0;
 Oy 1 MDKRTGWRGCHHVVEGLAGLEQLRRLBARLEHPOQR 35
 Db 154 LDBLPLAWEGGDNLRFARRRSFQARPATPPQGLR 189

RESULT 6

154361 SMX Protein - human
 N; Alternative names: escapes X-chromosome inactivation
 C; Species: Homo sapiens (man)
 C; Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 17-Mar-2000
 C; Accession: I154361; I38235
 R; Wu, J.; Ellison, J.; Salido, E.; Yen, P.; Mohandas, T.; Shapiro, L.J.
 Hum. Mol. Genet. 3, 153-160, 1994
 A; Title: Isolation and characterization of XE169, a novel human gene that escapes X-inactivation
 A; Reference number: I154361; MUID:9421434; PMID:8162017
 A; Accession: I154361
 A; Status: preliminary
 A; Molecule type: translated from GB/EMBL/DBJ
 A; Residues: 1-1560 <WU>
 A; Cross-references: GB:125270; NID:9457136; PID:9457137
 R; Agunik, A.I.; Mitchell, M.J.; Mattei, M.G.; Borzani, G.; Avner, P.A.; Lerner, J.L.; B hum. Mol. Genet. 3, 879-884, 1994
 A; Title: A novel X gene with a widely transcribed Y-linked homologue escapes X-inactivation
 C; Accession: I158235; MUID:95038739; PMID:7951230
 A; Accession: I138235
 A; Status: preliminary; translated from GB/EMBL/DBJ
 A; Molecule type: mRNA
 A; Residues: 280-341, 'Y', 343-344 <AGU>
 A; Cross-references: EMBL:229650; NID:9472994; PIDN:CAA82758.1; PID:9472995
 C; Genetics:
 A; Gene: GDB:SMX; XE169; DXS1272E
 A; Cross-references: GDB:9865442; OMIM:314690
 A; Map position: Xp11.22-Xp11.21
 C; Superfamily: human retinoblastoma binding protein 2
 Query Match 27.2%; Score 56.5; DB 2; Length 1560;
 Best Local Similarity 28.9%; Pred. No. 58; Mismatches 13; Conservative 7; MiMismatches 18; Indels 7; Gaps 1;
 Oy 1 MDEKTTGWRG-----GHVVEGLAGLEQLRRLBARLEHPOQR 38
 Db 1281 LITERAISWQRRARQALASDVTALLGRLABRQRLQASERPEEP 1325

RESULT 7
 E87711 heat shock protein HslV [imported] - Caulobacter crescentus
 C;Species: Caulobacter crescentus
 C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
 C;Accession: E87711
 R;Nieman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
 B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon.
 n., J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
 proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
 A;Title: Complete Genome Sequence of Caulobacter crescentus.
 A;Reference number: A87249; MUID:21173698; PMID:1259647
 A;Accession: E87711
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-188 <STO>
 A;Cross-references: GB:AB005673; NID:913425497; PIDN:AAK25689.1; GSPDB:GN00148
 C;Genetics:
 A;Gene: CC3727

Query Match 26.7%; Score 55.5; DB 2; Length 188;
 Best Local Similarity 43.3%; Pred. No. 8.2;
 Matches 13; Conservative 4; Mismatches 6; Indels 7; Gaps 1;
 Qy 10 GGAIVVEGLAGE-----LEQRLARLEHHP 32
 Db 52 GGKRVVAGFAGATAADFLLIERLEAKLQY 81

RESULT 8
 H87683 transcription regulator, AraC family [imported] - Caulobacter crescentus
 C;Species: Caulobacter crescentus
 C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
 C;Accession: H87683
 R;Nieman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
 B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon.
 n., J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
 proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
 A;Title: Complete Genome Sequence of Caulobacter crescentus.
 A;Reference number: A87249; MUID:21173698; PMID:1259647
 A;Accession: H87683
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-321 <STO>
 A;Cross-references: GB:AE005673; NID:913425234; PIDN:AAK25468.1; GSPDB:GN00148
 C;Genetics:
 A;Gene: CC3506

Query Match 26.0%; Score 54; DB 2; Length 321;
 Best Local Similarity 30.3%; Pred. No. 23; Mismatches 16; Indels 0; Gaps 0;
 Matches 10; Conservative 7; Mismatches 16; Indels 0; Gaps 0;
 Qy 2 DEKTTGWRGGHVVEGLACLEQRLARLEHHPQG 34
 Db 205 DPDPRDVWAAQQLIRGGAGTLDLARRIDILAPRG 237

RESULT 9
 E69444 chromosome segregation protein (smc1) homolog - Archaeoglobus fulgidus
 C;Species: Archaeoglobus fulgidus
 C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 02-Jun-2000
 C;Accession: E69444
 R;Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
 .; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.
 Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
 Nature 390, 364-370, 1997
 A;Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Arriach, P.; Kaine, B.P.; Sykes, S.
 Smith, H.O.; Woese, C.R.; Venter, J.C.
 A;Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeon
 A;Reference number: A69250; MUID:98049343; PMID:9389475

RESULT 10
 S68409 potassium channel Kv beta-3 chain - rat
 C;Species: Rattus norvegicus (Norway rat)
 C;Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 20-Jun-2000
 C;Accession: S68409; S72562
 R;Heinemann, S.H.; Rettig, J.; Wunder, F.; Pongs, O.
 PFS Lett. 377, 381-389, 1995
 A;Title: Molecular and functional characterization of a rat brain K(ν)-beta3 potassium c
 A;Reference number: S68409; MUID:96140552; PMID:8549760
 A;Experimental source: cortex
 C;Genetics:
 A;Gene: RChbeta3
 A;Molecule type: mRNA
 A;Residues: 1-104 <HLR>
 A;Cross-references: EMBL:X76723
 A;Experimental source: cortex
 R;Rettig, J.
 Submitted to the EMBL Data Library, December 1993
 A;Reference number: S72562
 A;Accession: S72562
 A;Molecule type: mRNA
 A;Residues: 1-245, 'P', 247-324, 'K', 326-404 <HLR>
 A;Cross-references: EMBL:X76723; NID:91246833; PIDN:CAA54141.1; PID:91246834
 A;Experimental source: cortex
 C;Genetics:
 A;Gene: RChbeta3
 C;Superfamily: fission yeast pyridoxine 4-dehydrogenase
 C;Keywords: voltage-gated ion channel

Query Match 25.7%; Score 53.5; DB 2; Length 404;
 Best Local Similarity 28.9%; Pred. No. 34; Mismatches 3; Indels 23; Gaps 2;
 Matches 13; Conservative 6; Mismatches 3; Indels 23; Gaps 2;
 Qy 4 KTTGWR-----GG-----HVVGVLAGEQLR 25
 Db 145 KSKGWRSSYYVITKFFWGGQATERGLSKHIGLQLSDRLQ 189

RESULT 11
 E87375 conserved hypothetical protein CCI017 [imported] - Caulobacter crescentus
 C;Species: Caulobacter crescentus
 C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
 C;Accession: E87375
 R;Nieman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
 B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon.
 n., J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
 proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
 A;Title: Complete Genome Sequence of Caulobacter crescentus.
 A;Reference number: A87249; MUID:21173698; PMID:1259647
 A;Accession: E87375
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-554 <STO>
 A;Cross-references: GB:AB005673; NID:913422309; PIDN:AAK23001.1; GSPDB:GN00148
 C;Genetics:
 A;Gene: CC1017

Query Match 25.7%; Score 53.5; DB 2; Length 554;
 Best Local Similarity 35.9%; Pred. No. 47; Mismatches 14; Conservative 4; Indels 5; Gaps 1; Matches 14; Mismatches 16; Indels 5; Gaps 1;

Qy 5 TPGWRSGHVWE----GLAGEFLQLRLEHPPQREP 38
 Db 45 STEWSAGVYTDVNTFGYYGELNPLRCLPLLTUVGRHAP 83

RESULT 12

C75568 hypothetical protein - *Deinococcus radiodurans* (strain R1)
 C;Species: *Deinococcus radiodurans*
 C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000
 C;Accession: C75568
 C;Text: Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1.
 R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; Ma, S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
 A;Title: Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1.
 A;Reference number: A75250; MUID:20036896; PMID:10567256
 A;Accession: C75568
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-261 <WHI>
 A;Cross-references: GB:AE001867; GB:AE000513; NID:96457693; PIDN:AAF09637.1; PID:9645770
 A;Experimental source: strain R1
 C;Genetics:
 A;Gene: DR0043
 A;Map position: 1

RESULT 13

S75456 protein plsx - *Synechocystis* sp. (strain PCC 6803)
 N;Alternative names: protein slyr510
 C;Species: *Synechocystis* sp.
 A;Variety: PCC 6803
 C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 28-Jul-2003
 C;Accession: S75456
 C;Text: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* sp. A;Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* sp.
 A;Reference number: S74322; MUID:97061201; PMID:8905231
 A;Accession: S75456
 A;Status: nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-348 <KAN>
 A;Cross-references: EMBL:D90911; GB:AB001339; NID:91653083; PIDN:BA18017.1; PID:9165310
 C;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
 C;Genetics:
 A;Gene: plsx
 C;Superfamily: phospholipid biosynthesis protein, Plsx type
 Query Match 25.5%; Score 53; DB 1; Length 348;
 Best Local Similarity 34.3%; Pred. No. 33; Mismatches 12; Conservative 9; Indels 4; Gaps 2; Matches 12; Mismatches 10; Indels 4; Gaps 2;

Qy 2 DEKTTGHRG-GHUVVEGLAGELFQLRLEHPPQ 34
 Db 255 BELLPRGRGKGKGAII-LAPPNKRKIQKVDIHEHG 287

RESULT 14

H70668 probable polyketide synthase - *Mycobacterium tuberculosis* (strain H37RV)
 H70668 C;Species: *Mycobacterium tuberculosis*
 C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
 C;Accession: H70668; S73074
 C;Text: Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome : A;Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome : A;Reference number: A70500; MUID:98295987; PMID:9634230
 A;Accession: H70668
 A;Status: preliminary; nucleic acid sequence not shown
 A;Molecule type: DNA
 A;Description: Mycobacterium tuberculosis cosmid tbc2.
 A;Reference number: S7053
 A;Accession: S73074
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-273, 275-496 <SMI>
 A;Cross-references: EMBL:U00024; NID:9560506; PIDN:AAA50929.1; PID:9560508
 C;Genetics:
 A;Gene: pks15
 A;Start codon: GTG
 F;Superfamily: 3'-oxoacyl-[acyl-carrier-protein] synthase I; 3'-oxoacyl-[acyl-carrier-protein] synthase I homology <QAS2>
 F;68-467/Domain: 3'-oxoacyl-[acyl-carrier-protein] synthase I homology <QAS2>
 Query Match 25.5%; Score 53; DB 2; Length 496;
 Best Local Similarity 40.7%; Pred. No. 49; Mismatches 11; Conservative 6; Indels 10; Gaps 0; Matches 11; Conservative 6; Mismatches 10; Indels 0; Gaps 0;
 Qy 12 HVVEGLAGELFQLRLEHPPQREP 38
 Db 22 HYLKKVAVELDETRARLREYERATEP 48

RESULT 15

S06954 intermediate filament protein B - common roundworm
 C;Species: *Ascaris lumbricoides* (common roundworm)
 C;Accession: S06954
 C;Text: Cyttoplasmic intermediate filament proteins of invertebrates are closer to nucleic acid of a nematode.
 A;Reference number: S06954; MUID:90059912; PMID:2580097
 A;Accession: S06954
 A;Molecule type: protein
 A;Residues: 1-589 <WEB>
 A;Cross-references: EMBL:U00024; NID:91653083; PIDN:BA18017.1; PID:9165310
 C;Superfamily: intermediate filament protein Av71
 C;Keywords: coiled coil
 Query Match 25.5%; Score 53; DB 2; Length 589;
 Best Local Similarity 48.0%; Pred. No. 59; Mismatches 12; Conservative 2; Indels 2; Gaps 1; Matches 12; Mismatches 9; Indels 2; Gaps 1;

Qy 3 EKTGWRGRGHTVVEGLAGELFQLR 27
 Db 182 EKINQWQ-HAIEDAQELEMRLAR 204

Search completed: July 15, 2004, 20:36:21
 Job time : 11.2824 secs

Copyright (c) 1993 - 2004 Compugen Ltd.	GenCore version 5.1.6	28	52	25.0	261	4	US-09-252-991A-28545	Sequence 28545, A
Run on:	July 15, 2004, 20:32:20 ; Search time 12.9647 Seconds	29	52	25.0	423	2	US-08-715-554-2	Sequence 2, Appli
OM protein - protein search, using sw model		30	52	25.0	423	2	US-08-581-148-2	Sequence 2, Appli
Scoring table:	BLOSUM62	31	52	25.0	499	4	US-09-252-991A-27221	Sequence 27221, A
Post-processing: Minimum Match 0%	Maximum DB seq length: 0	32	52	25.0	741	4	US-09-252-991A-27062	Sequence 27062, A
Perfect score: 208	Sequence: 1 MDEKITGWRGGHVVVEGLAGLEQLRARLBRHHPQGQREP 38	33	51	24.8	174	4	US-09-252-991A-31496	Sequence 31496, A
Scoring table: Gapop 10.0 , Gapext 0.5		34	51.5	24.8	248	4	US-09-252-991A-30679	Sequence 1, Appli
Total number of hits satisfying chosen parameters: 389414		35	51.5	24.8	329	2	US-08-606-143-1	Sequence 3, Appli
Minimum DB seq length: 0		36	51.5	24.8	329	2	US-08-606-143-3	Sequence 31542, A
Maximum DB seq length: 2000000000		37	51	24.5	166	4	US-09-328-352-5059	Sequence 5059, AP
Post-processing: Maximum Match 100%	Listing first 45 summaries	38	51	24.5	179	4	US-09-252-991A-24661	Sequence 24661, A
Database : Issued Patents AA:*		39	51	24.5	248	4	US-09-252-991A-30679	Sequence 30679, A
1: /cgn2_6/ptodata/2/iaa/5A-COMB.pep:*		40	51	24.5	293	4	US-09-252-991A-27745	Sequence 27745, A
2: /cgn2_6/ptodata/2/iaa/5B-COMB.pep:*		41	51	24.5	324	4	US-09-252-991A-28815	Sequence 28815, A
3: /cgn2_6/ptodata/2/iaa/6A-COMB.pep:*		42	51	24.5	427	4	US-09-489-039A-0737	Sequence 10737, A
4: /cgn2_6/ptodata/2/iaa/6B-COMB.pep:*		43	51	24.5	428	4	US-09-252-991A-20199	Sequence 20199, A
5: /cgn2_6/ptodata/2/iaa/PCTUS.COMB.pep:*		44	51	24.5	430	1	US-08-631-607-7	Sequence 7, Appli
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*		45	51	24.5	430	4	US-09-098-358B-7	Sequence 7, Appli
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.	SUMMARIES							
	RESULT 1							
	; Sequence 91, Application US/09413814							
	; Patent No. 6225064							
	; GENERAL INFORMATION:							
	; APPLICANT: Gesellschaft fuer Biotechnologische Forschung mbH							
	; APPLICANT: Bristol-Myers Squibb, Co.							
	; APPLICANT: Beyer, Stefan							
	; APPLICANT: Bloecker, Helmut							
	; APPLICANT: Brandt, Petra							
	; APPLICANT: Cino, Paul M							
	; APPLICANT: Dougherty, Brian A							
	; APPLICANT: Goldberg, Steven L							
	; APPLICANT: Hofle, Gerhard							
	; APPLICANT: Mueller, Joachim							
	; APPLICANT: Reichenbach, Hans							
	; TITLE OF INVENTION: DNA sequences for enzymatic synthesis of polyketide or							
	; FILE REFERENCE: PCT/US 99/23535							
	; CURRENT APPLICATION NUMBER: US/09/413, 814							
	; EARLIER APPLICATION NUMBER: DE 198 46 493 .2							
	; EARLIER FILING DATE: 1998-10-09							
	; NUMBER OF SEQ ID NOS: 107							
	; SEQ ID NO 91							
	; LENGTH: 543							
	; TYPE: PRT							
	; ORGANISM: <i>Sorangium cellulosum</i>							
	US-09-413-814-91							
	Query Match	30.3*						
	Best Local Similarity	35.0*						
	Matches	13;						
	Sequence 1, Conservative	6;						
	Sequence 2, Applied	6;						
	Sequence 3, Applied	6;						
	Sequence 4, Applied	6;						
	Sequence 5, Applied	6;						
	Sequence 6, Applied	6;						
	Sequence 7, Applied	6;						
	Sequence 8, Applied	6;						
	Sequence 9, Applied	6;						
	Sequence 10, Applied	6;						
	Sequence 11, Applied	6;						
	Sequence 12, Applied	6;						
	Sequence 13, Applied	6;						
	Sequence 14, Applied	6;						
	Sequence 15, Applied	6;						
	Sequence 16, Applied	6;						
	Sequence 17, Applied	6;						
	Sequence 18, Applied	6;						
	Sequence 19, Applied	6;						
	Sequence 20, Applied	6;						
	Sequence 21, Applied	6;						
	Sequence 22, Applied	6;						
	Sequence 23, Applied	6;						
	Sequence 24, Applied	6;						
	Sequence 25, Applied	6;						
	Sequence 26, Applied	6;						
	Sequence 27, Applied	6;						
	Sequence 28, Applied	6;						
	Sequence 29, Applied	6;						
	Sequence 30, Applied	6;						
	Sequence 31, Applied	6;						
	Sequence 32, Applied	6;						
	Sequence 33, Applied	6;						
	Sequence 34, Applied	6;						
	Sequence 35, Applied	6;						
	Sequence 36, Applied	6;						
	Sequence 37, Applied	6;						
	Sequence 38, Applied	6;						
	Sequence 39, Applied	6;						
	Sequence 40, Applied	6;						
	Sequence 41, Applied	6;						
	Sequence 42, Applied	6;						
	Sequence 43, Applied	6;						
	Sequence 44, Applied	6;						
	Sequence 45, Applied	6;						
	Sequence 46, Applied	6;						
	Sequence 47, Applied	6;						
	Sequence 48, Applied	6;						
	Sequence 49, Applied	6;						
	Sequence 50, Applied	6;						
	Sequence 51, Applied	6;						
	Sequence 52, Applied	6;						
	Sequence 53, Applied	6;						
	Sequence 54, Applied	6;						
	Sequence 55, Applied	6;						
	Sequence 56, Applied	6;						
	Sequence 57, Applied	6;						
	Sequence 58, Applied	6;						
	Sequence 59, Applied	6;						
	Sequence 60, Applied	6;						
	Sequence 61, Applied	6;						
	Sequence 62, Applied	6;						
	Sequence 63, Applied	6;						
	Sequence 64, Applied	6;						
	Sequence 65, Applied	6;						
	Sequence 66, Applied	6;						
	Sequence 67, Applied	6;						
	Sequence 68, Applied	6;						
	Sequence 69, Applied	6;						
	Sequence 70, Applied	6;						
	Sequence 71, Applied	6;						
	Sequence 72, Applied	6;						
	Sequence 73, Applied	6;						
	Sequence 74, Applied	6;						
	Sequence 75, Applied	6;						
	Sequence 76, Applied	6;						
	Sequence 77, Applied	6;						
	Sequence 78, Applied	6;						
	Sequence 79, Applied	6;						
	Sequence 80, Applied	6;						
	Sequence 81, Applied	6;						
	Sequence 82, Applied	6;						
	Sequence 83, Applied	6;						
	Sequence 84, Applied	6;						
	Sequence 85, Applied	6;						
	Sequence 86, Applied	6;						
	Sequence 87, Applied	6;						
	Sequence 88, Applied	6;						
	Sequence 89, Applied	6;						
	Sequence 90, Applied	6;						
	Sequence 91, Applied	6;						
	Sequence 92, Applied	6;						
	Sequence 93, Applied	6;						
	Sequence 94, Applied	6;						
	Sequence 95, Applied	6;						
	Sequence 96, Applied	6;						
	Sequence 97, Applied	6;						
	Sequence 98, Applied	6;						
	Sequence 99, Applied	6;						
	Sequence 100, Applied	6;						
	Sequence 101, Applied	6;						
	Sequence 102, Applied	6;						
	Sequence 103, Applied	6;						
	Sequence 104, Applied	6;						
	Sequence 105, Applied	6;						
	Sequence 106, Applied	6;						
	Sequence 107, Applied	6;						
	Sequence 108, Applied	6;						
	Sequence 109, Applied	6;						
	Sequence 110, Applied	6;						
	Sequence 111, Applied	6;						
	Sequence 112, Applied	6;						
	Sequence 113, Applied	6;						
	Sequence 114, Applied	6;						
	Sequence 115, Applied	6;						
	Sequence 116, Applied	6;						
	Sequence 117, Applied	6;						
	Sequence 118, Applied	6;						
	Sequence 119, Applied	6;						
	Sequence 120, Applied	6;						
	Sequence 121, Applied	6;						
	Sequence 122, Applied	6;						
	Sequence 123, Applied	6;						
	Sequence 124, Applied	6;						
	Sequence 125, Applied	6;						
	Sequence 126, Applied	6;						
	Sequence 127, Applied	6;						
	Sequence 128, Applied	6;						
	Sequence 129, Applied	6;						
	Sequence 130, Applied	6;						
	Sequence 131, Applied	6;						
	Sequence 132, Applied	6;						
	Sequence 133, Applied	6;						
	Sequence 134, Applied	6;						
	Sequence 135, Applied	6;						
	Sequence 136, Applied	6;						
	Sequence 137, Applied	6;						
	Sequence 138, Applied	6;						
	Sequence 139, Applied	6;						
	Sequence 140, Applied	6;						
	Sequence 141, Applied	6;						
	Sequence 142, Applied	6;						
	Sequence 143, Applied	6;						
	Sequence 144, Applied	6;						
	Sequence 145, Applied	6;						
	Sequence 146, Applied	6;						
	Sequence 147, Applied	6;						
	Sequence 148, Applied	6;						
	Sequence 149, Applied	6;						
	Sequence 150, Applied	6;						
	Sequence 151, Applied	6;						
	Sequence 152,							

; APPLICANT: Cino, Paul M
 ; APPLICANT: Dougherty, Brian A
 ; APPLICANT: Goldberg, Steven L
 ; APPLICANT: Horle, Gerhard
 ; APPLICANT: Mueller, Joachim
 ; APPLICANT: Reichenbach, Hans
 ; TITLE OF INVENTION: DNA Sequences for enzymatic synthesis of polyketide or
 ; TITLE OF INVENTION: heteropolyketide compounds
 ; FILE REFERENCE: PCT/US 99/25353
 ; CURRENT APPLICATION NUMBER: US/09/413, 814
 ; CURRENT FILING DATE: 1999-10-07
 ; EARLIER APPLICATION NUMBER: DE 198 46 493.2
 ; EARLIER FILING DATE: 1998-10-09
 ; NUMBER OF SEQ ID NOS: 107
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 81
 ; LENGTH: 544
 ; TYPE: PRT
 ; ORGANISM: Sorangium cellulosum
 ; US-09-413-814-81

 RESULT 3
 Query Match 30.3%; Score 63; DB 3; Length 544;
 Best Local Similarity 35.1%; Pred. No. 0.97; Matches 13; Conservative 6; Mismatches 10; Indels 8; Gaps 1;
 Matches 13; Conservative 6; Mismatches 10; Indels 8; Gaps 1;
 Qy 8 WRGGHVVVEG-----LAGEBQLARLHHHQGR 36
 Qy 8 WRGGHVVVEG-----LAGEBQLARLHHHQGR 36
 Db 236 WLSGEVLEGRLRYWTTKLSGALLRARRVFDHEPAGR 272
 Db 112 WRGGRAPEGPQGRHRRLRRHHRPQGR 140

 RESULT 4
 US-09-252-991A-29830
 Sequence 29830. Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252, 991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIORITY APPLICATION NUMBER: US 60/074, 788
 ; PRIORITY FILING DATE: 1998-02-18
 ; PRIORITY APPLICATION NUMBER: US 60/094, 190
 ; PRIORITY FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 29830
 ; LENGTH: 543
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 ; US-09-252-991A-29830

 Query Match 28.6%; Score 59.5; DB 4; Length 543;
 Best Local Similarity 42.9%; Pred. No. 3; Matches 15; Conservative 3; Mismatches 14; Indels 3; Gaps 2;
 Matches 15; Conservative 3; Mismatches 14; Indels 3; Gaps 2;
 Qy 7 GRGGHVVVEGLGG-----ELBQLARLHHHQ-3GREP 38
 Qy 7 GRGGHVVVEGLGG-----ELBQLARLHHHQ-3GREP 38
 Db 183 GOSAGHAFQGLADLVDLQDLYVAQRHHFRPDHHP 217

 RESULT 5
 US-08-248-839C-6
 Sequence 6. Application US/08248839C
 ; Patent No. 5843702
 ; GENERAL INFORMATION:
 ; APPLICANT: McConnell, David
 ; APPLICANT: Devine, Kevin
 ; APPLICANT: O'Kane, Charles
 ; TITLE OF INVENTION: A Gene Expression System
 ; NUMBER OF SEQUENCES: 185
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESS: No. 5843702 disk of No. 5843702th America, Inc.
 ; STREET: 405 Lexington Avenue
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: USA
 ; ZIP: 10174-6401
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSEQ for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/248, 839C
 ; FILING DATE: 25-MAY-1994
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Gregg Valeta A.
 ; REGISTRATION NUMBER: 35,127
 ; REFERENCE/DOCKET NUMBER: 3614.214-US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 212-878-9655
 ; TELEFAX: 212-867-0123
 ; INFORMATION FOR SEQ ID NO: 6:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 113 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: Protein
 ; US-08-248-839C-6

 Query Match 27.9%; Score 58; DB 2;
 Best Local Similarity 40.0%; Pred. No. 0.81; Matches 12; Conservative 4; Mismatches 14; Indels 0; Gaps 0;
 Matches 12; Conservative 4; Mismatches 14; Indels 0; Gaps 0;
 Qy 9 RGHHVVEGLGG-----ELBQLARLHHHQ-3GREP 38
 Qy 9 RGHHVVEGLGG-----ELBQLARLHHHQ-3GREP 38
 Db 10 RGKRTQEEIAGHIGVSRARYSHYENGSEP 39
 Db 10 RGKRTQEEIAGHIGVSRARYSHYENGSEP 39

 RESULT 6
 US-03-252-991A-22131
 ; Sequence 22131. Application US/09252991A
 ; Patent No. 6551795

GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196-136
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 22131
; LENGTH: 628
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-22131

Query Match 27.9%; Score 59; DB 4; Length 628;
Best Local Similarity 44.4%; Pred. No. 5.8; Mismatches 16; Conservative 1; Indels 11; Gaps 8; Gaps 1;
Db 126 RGHHAVEVGLAVGGGLAVELGAVPGRGADHPVGPR 36

RESULT 7
US-09-134-000C-5025
; Sequence 5025, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEAR ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5025
; LENGTH: 239
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
; US-09-134-000C-5025

Query Match 27.6%; Score 57.5; DB 4; Length 239;
Best Local Similarity 52.2%; Pred. No. 2.3; Mismatches 12; Conservative 3; Mismatches 7; Indels 1; Gaps 1;
Db 8 WRGGHVVVEGLAGELBQLRLEH 30

RESULT 8
US-09-252-991A-19549
; Sequence 19549, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEAR ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196-136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 2
; LENGTH: 1388
; TYPE: PRT
; ORGANISM: Human
; US-09-252-991A-19549

Query Match 27.4%; Score 57; DB 4; Length 356;
Best Local Similarity 45.5%; Pred. No. 4.2; Mismatches 15; Conservative 3; Mismatches 9; Indels 6; Gaps 2;
Db 7 GHRGGHVVVEGLAGELBQLRLEHHRHQ-GORE 37

RESULT 9
US-09-252-991A-29119
; Sequence 29119, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEAR ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196-136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 29119
; LENGTH: 904
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-29119

Query Match 27.4%; Score 57; DB 4; Length 904;
Best Local Similarity 36.1%; Pred. No. 12; Mismatches 13; Conservative 3; Mismatches 20; Indels 0; Gaps 0;
Db 1 1 MDRKTTGWRGGHVVVEGLAGELBQLRLEHHRQGR 36

RESULT 10
US-09-572-191-2
; Sequence 2, Application US/09572191
; Patent No. 6355466
; GENERAL INFORMATION:
; APPLICANT: Belaud, Christophe
; APPLICANT: Sakowicz, Roman
; APPLICANT: Wood, Kenneth
; TITLE OF INVENTION: No. 6355466el motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1017
; CURRENT APPLICATION NUMBER: US/09/572,191
; CURRENT FILING DATE: 2000-05-17
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1388
; TYPE: PRT
; ORGANISM: Human
; US-09-572-191-2

Query Match 26.7%; Score 55.5; DB 4; Length 1388;
Best Local Similarity 37.5%; Pred. No. 32; Mismatches 12; Conservative 6; Mismatches 7; Indels 7; Gaps 1;

RESULT 11
 US-09-723-262-2
 ; Sequence 2, Application US/09723262
 ; Patent No. 6379912
 ; GENERAL INFORMATION:
 ; APPLICANT: Beraud, Christophe
 ; APPLICANT: Sakowicz, Roman
 ; TITLE OF INVENTION: No. 6379912el motor proteins and methods for
 ; TITLE OF INVENTION: their use
 ; FILE REFERENCE: 1017
 ; CURRENT APPLICATION NUMBER: US/09/723, 262
 ; CURRENT FILING DATE: 2000-11-27
 ; PRIORITY APPLICATION NUMBER: US 09/572, 191
 ; PRIORITY FILING DATE: 2000-05-17
 ; NUMBER OF SEQ ID NOS: 6
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO 2
 ; LENGTH: 1388
 ; TYPE: PRT
 ; ORGANISM: Human
 ; US-09-723-262-2

Query Match 26.7%; Score 55.5; DB 4; Length 1388;
 Best Local Similarity 37.5%; Pred. No. 32; Mismatches 7; Indels 7; Gaps 1;
 Matches 12; Conservative 6; Mismatches 7; Indels 7; Gaps 1;

QY 9 RSGHVV-----GLAGLEQQLRQLRHHPO 33
 Db 478 RGGFLPPEQQDRILSELNEIQTLREQIEHPPR 509

RESULT 12
 US-09-723-219-2
 ; Sequence 2, Application US/09723219
 ; Patent No. 6391613
 ; GENERAL INFORMATION:
 ; APPLICANT: Beraud, Christophe
 ; APPLICANT: Sakowicz, Roman
 ; APPLICANT: Wood, Kenneth
 ; TITLE OF INVENTION: No. 6391613el motor proteins and methods for
 ; TITLE OF INVENTION: their use
 ; FILE REFERENCE: 1017
 ; CURRENT APPLICATION NUMBER: US/09/723, 219
 ; CURRENT FILING DATE: 2000-11-27
 ; PRIORITY APPLICATION NUMBER: US 09/572, 191
 ; PRIORITY FILING DATE: 2000-05-17
 ; NUMBER OF SEQ ID NOS: 6
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO 2
 ; LENGTH: 1388
 ; TYPE: PRT
 ; ORGANISM: Human
 ; US-09-723-219-2

Query Match 26.0%; Score 54; DB 4; Length 185;
 Best Local Similarity 40.5%; Pred. No. 52; Mismatches 4; Indels 4; Gaps 2;
 Matches 15; Conservative 4; Mismatches 14; Indels 4; Gaps 2;

QY 2 DEKTTGWRGGHVVFGAGLEQQLRQLRHH---POQQ 35
 Db 98 DEMEBGWEGCLSVPGLRGAVSRHR-RIRYQGDPOQQ 133

RESULT 14
 US-09-252-991A-24870
 ; Sequence 24870, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; TITLE OF INVENTION: ARROGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107195-136
 ; CURRENT APPLICATION NUMBER: US/09/252, 991A
 ; PRIORITY APPLICATION NUMBER: US 60/094, 190
 ; PRIORITY FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 19647
 ; LENGTH: 185
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 ; US-09-252-991A-24870

Query Match 25.7%; Score 53.5; DB 4; Length 266;
 Best Local Similarity 34.2%; Pred. No. 93; Mismatches 6; Indels 15; Gaps 1;
 Matches 13; Conservative 4; Mismatches 6; Indels 15; Gaps 1;

QY 15 EJLAGLEQQLRQLRHH-----QGORE 37
 Db 112 QGVEPELLEQQLRQAVDHHQHQLPCREHQAQLVQGHE 149

RESULT 15
 US-09-252-991A-19069
 ; Sequence 19069, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; TITLE OF INVENTION: ARROGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107195-136
 ; CURRENT APPLICATION NUMBER: US/09/252, 991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIORITY APPLICATION NUMBER: US 60/094, 190
 ; PRIORITY FILING DATE: 1998-02-18

PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 19069
LENGTH: 692
TYPE: PRT
ORGANISM: *Pseudomonas aeruginosa*
US-09-252-991A-19069

Query Match 25.7%; Score 53.5; DB 4; Length 692;
Best Local Similarity 42.9%; Pred. No. 28;
Matches 12; Conservative 4; Mismatches 9;
Indels 3; Gaps 1;
Matches 12; Conservative 4; Mismatches 9;
Indels 3; Gaps 1;

QY	6	TGRRGGHVV EG AG EL EQ RL BR HH PQ	33
Db	287	TGRRGEV IE GSV GL RL--AERHPR	311

Search completed: July 15, 2004, 20:37:00
Job time : 12.9647 secs

This Page Blank (uspto)

This Page Blank (uspto)